Economics and Human Biology 47 (2022) 101194

Contents lists available at ScienceDirect

Economics and Human Biology

journal homepage: www.elsevier.com/locate/ehb

Education and COVID-19 excess mortality

Piera Bello [a], Lorenzo Rocco [b,*]

[a] University of Bergamo, Italy, and ZEW, Germany
[b] University of Padova, Italy, and IZA, Germany

ARTICLE INFO

JEL classification in this article:
I14
I18
I26
R00

Keywords:
COVID-19
Education
Excess mortality
Municipality
Parallel trend
Fake news
Public health communication

ABSTRACT

We study the role of education during the COVID-19 epidemic in Italy. We compare excess mortality in 2020 and 2021 compared to the pre-pandemic mortality between municipalities with different shares of educated residents. We find that education initially played a strong protective role, which however quickly faded out. After pondering several alternative explanations, we tentatively interpret this finding as the outcome of the interplay between education, information and public health communication, whose availability and coherence varied along the epidemic.

1. Introduction

COVID-19 mainly spreads through close contact, from an infected person's mouth or nose in small liquid particles when they cough, sneeze, speak, sing or breathe (WHO, 2020a). It is thus possible to reduce the risk of getting infected and infecting, by keeping social distancing, using face masks and adopting specific hygiene practices. These are the strategies that were eventually recommended by the health authorities and experts all around the world (Cheng et al., 2021), and that also encountered quite some resistance in the public because they limit social interactions, the freedom of movement and face the inertia of long-established and deeply interiorised social behaviours (Ornaghi and Tonin, 2018). As everyone differently weights the benefits and the costs of these strategies, their adoption was scattered in the population. Galasso et al. (2020), and Desmet and Wacziarg (2021), among others, show that the adoption of protective strategies depended on gender, age, occupation, income, trust in the government and beliefs. Bartscher et al. (2021) (among others) suggest that adoption was stronger in regions richer of social capital.

A distinctive factor that plausibly affects adoption is education. We hypothesise that education enables a correct assessment of the effectiveness of protective strategies (Reyna et al., 2009), by influencing people ability to access and process information, judge the risk of infection and the credibility of information providers, discriminate between evidence-based and fake news (Freeman et al., 2020). Our paper is the first study, to the best of our knowledge, that investigates if and to what extent education played an autonomous role in determining the burden of the COVID-19 pandemic.

We focus on Italy and we relate municipality average education with municipality excess mortality, i.e. the increase in the mortality rate among the population aged over 60 between 2020 (or 2021) and the pre-COVID period. Excess mortality captures the mortality both directly and indirectly related to COVID-19, the latter being mediated by possible congestion of the healthcare system at the peak of the epidemic, delays in programmed treatments, surgery and screening, and provides an accurate account of the intensity of the pandemic (Sanmarchi et al., 2021; Beaney et al., 2020). A distinctive advantage of using excess mortality as an indicator rather than the incidence of COVID-19, hospitalizations or COVID-19 mortality is that it is independent of any count of virus cases, which can be largely underestimated (Wu et al., 2020; Irons and Raftery, 2021), and it is robust to mis-classification of the cause of death. The lag between infection and death is difficult to estimate, but it is safe to conclude that in most cases it is less than 20 days (Wiliński et al., 2022; Marschner, 2021, Ward and Johnsen, 2021), implying that the excess

* Correspondence to: Department of Economics and Management, University of Padova, Via del Santo, 33 - 35123 Padova, Italy.
E-mail addresses: piera.bello@unibg.it (P. Bello), lorenzo.rocco@unipd.it (L. Rocco).

https://doi.org/10.1016/j.ehb.2022.101194
Received 29 March 2022; Received in revised form 14 October 2022; Accepted 21 October 2022
Available online 25 October 2022
1570-677X/

mortality recorded in a period of time of two or three months is largely correlated with the spread of the virus in the same period.

We consider mortality in the population aged 60 and over, because this is the age group that almost exclusively suffered the worst consequences of COVID-19, and we split the epidemic years 2020 and 2021 into periods: January to February; March to May; June to September; and October to December. Both in 2020 and 2021 a COVID-19 wave mounted between March and May, while in both years circulation of the virus was much lower in the summer period June-September. January-February 2020 was the pre-pandemic period, while January-February 2021 corresponds to the decline of the winter wave that raised between October-December 2020. October-December 2021 corresponds to the early stages of the Omicron wave, but it is not covered by our data. Overall, we analyse seven periods of, on average, three months each, over two years. We also distinguish between North and Centre-South Italy, because the timing of the first epidemic outbreak differed in the two areas. Northern Italy experienced the first massive circulation of the virus between March and May 2020 (the first wave) while the Centre-South was reached by the pandemic only later, between October 2020 and February 2021 (the second wave).

For each area and each period/year, we test whether the COVID-19-induced excess mortality is larger or smaller in municipalities with higher education attainment than in municipalities with lower attainment, other characteristics being equal.[1] There are two empirical issues to be addressed in our analysis. First, we need to verify that mortality trends *before* the pandemic are parallel between municipalities with different levels of education, so that the correlation between education and excess mortality *during* the pandemic can be attributed to the specific role that education played against COVID-19. Second, we need to isolate the effect of education from that of possible confounders, which might also have played a role in the pandemic. To this purpose, we control for an extensive list of measures of economic affluence, mobility, urbanization, the degree of teleworkability among the employed, social capital, and demographics among others.

We find that in Northern Italy, during the Spring wave of 2020, education played a significant protective role. Excess mortality was much smaller in the municipalities with higher levels of education than among their less educated counterparts. Specifically, a 10-percentage point increase in the share of residents with at least secondary education was associated with 0.426 deaths per 1000 inhabitants lower excess mortality, which correspond to about 24% of the average excess mortality registered in the North between March and May 2020 compared to pre-pandemic years. During the Winter wave of 2020, which was the first extending to the Centre-South, we do not find any significant effect related to education in the North and some evidence of a moderate protective effect in the Centre-South. No effect is detected in the Spring wave of 2021 in both areas. Overall, we find evidence of a protective effect of education when an area is first reached by COVID-19 but not thereafter.

Interpreting why the effect of education varies as the pandemic progresses is difficult, especially for the lack of timely and geographically detailed data on individual behaviour, information and perceptions. Our large set of controls allows to rule out several spurious effects of education, those due to the correlation between education and local characteristics, such as urbanization, age, employment rates, availability of health care services and social capital, among others. After discussing the plausibility of some alternatives, our preferred interpretation points to the interplay between education and the information set available to individuals, which includes official public health communication, information provided by non-specialized media and word-of-mouth. Several papers document indeed the surge of the so called infodemic, which is the wave of fake news and unscientific claims that accompanied the pandemic (WHO, 2021; Zarocostas, 2020; Gallotti et al., 2020; Loomba, 2021; Yang et al., 2021).[24] Our conjecture is that the effect of education belongs to the cognitive domain and its intensity depends on the quality of the information available to individuals. When messages are contradictory, as it was particularly the case in the early phases of the pandemic, the more educated are better able to discriminate information and hence protect more effectively than the less educated. Differently, when information is more reliable and univocal, and/or when people have learnt from their or others' experience, as it was the case later in the pandemic, individuals are able to adopt proper protective strategies independently of their education.

The rest of the paper is organised as follows. Section 2 is devoted to the review of the relevant literature. Section 3 documents the phases of the pandemic in Italy and the evolution of the guidelines from March 2020 onwards. Section 4 describes our data, while the empirical strategy is discussed in Section 5. Results are collected in Section 6. Possible interpretations are discussed in Section 7 and a summary of the findings and policy implications follow in the Conclusions. Two Appendixes collect details about the model specification and the progression the quality of information over our study period.

2. Relevant literature

Evaluating the role of education in explaining COVID-19 mortality is part of the ampler problem of assessing the causal role of education on health. Extending early studies from Adams (2002) and Arendt (2005), Cutler and Lleras-Muney (2010) investigate the education gradient, which is the link between education and health behaviour and is responsible for huge differences in life expectancy. They conclude that a key component of the gradient is that education rises cognition and, in turn, improves behaviour. Conti et al. (2010) document that education has a strong causal effect on health behaviour, physical and mental health, although the extent of the effect is heterogeneous. More recently, Hong et al. (2020) estimate the effect of college vs non-college education in the US and conclude that education improves health at a later age and sizeably extends life expectancy among males.

Little attention has been devoted to the role of education in the COVID-19 pandemic, despite the plausible influence that education can have in the adoption of protective strategies. We have been able to find only a few papers that study the effect of education, and also in these cases, education is not the paper's main focus. Indeed, Charoenwong et al. (2020) use data from Facebook in the US by county and estimate a DID model where they exploit the scattered introduction of mobility restrictions in the local areas at the beginning of the pandemic. They find that counties with a higher proportion of college graduates are more responsive and increase more their social distancing. They also find that having Facebook connections with less educated individuals decreases compliance. Armillei et al. (2021) analyse the central/periphery gradient in the excess mortality caused by COVID-19 among Italian municipalities and attribute part of this gradient to differential education levels. Evidence from previous epidemics reviewed by Bish and Michie (2010) indicates that compliance is positively associated with education. However, Wright et al. (2021), who exploit the COVID-19 Social Study, a longitudinal study that follows a panel of about 50,000

[1] Estimating the effect of education on excess mortality within each area and period, without exploiting any comparison between areas and periods, has two advantages. First, it neutralizes seasonal effects (Healy, 2003), such as those related to the seasonal flu and the so-called excess winter mortality (Rolfes et al., 2018; Crighton et al., 2007). Second, it addresses the point recently made by Callaway and Li (2021) and Bisin and Moro (2020) that the highly non-linear dynamics of the pandemic bias the estimates obtained from strategies that exploit variations in the timing and evolution of the epidemic.

[24] For 2020 and 2021, the minor differences between the point estimates in Table A1 and those in Table 2 and 3 are due to the different reference period used in the two models, namely year 2012 in model (2) and the average between 2012 and 2019 in model (1).

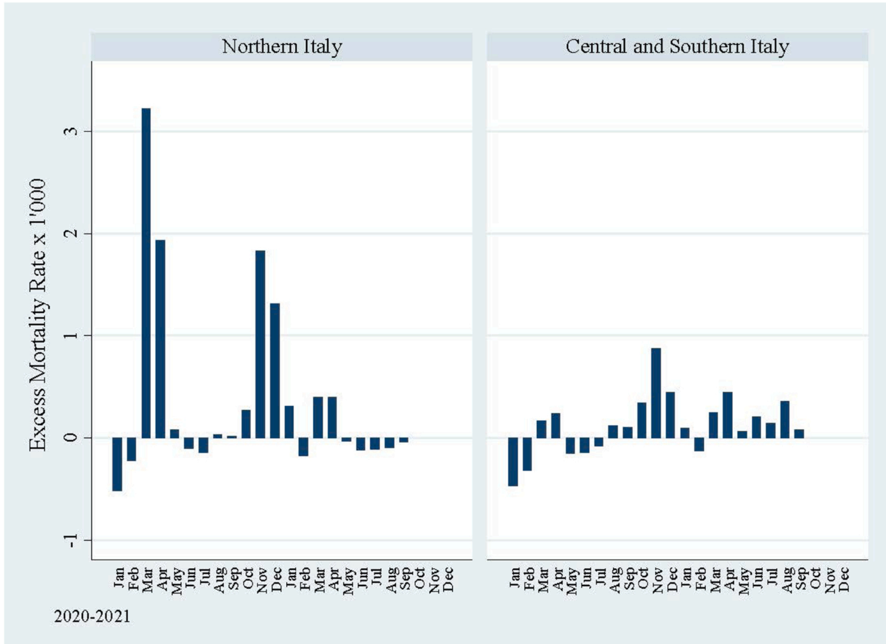


Fig. 1. Average excess mortality in 2020 and 2021 compared to the average 2012–2019, by month. Note: The figure shows the average excess mortality in Northern and in Central-Southern Italian municipalities for each month of 2020 and 2021. Excess mortality is the difference between the average mortality rate in 2020 for the age group 60 + and its 2012–2019 average.
Source: ISTAT.

Britons during the pandemic, do not find evidence that compliance is related to education. Overall, the scant evidence about the effect of education is mixed. Charoenwong et al. (2020) and Armillei et al. (2021) suggest a protective role, while others argue the opposite (Wright et al., 2021; Nivette et al., 2021; Galasso et al., 2020). We contribute to this debate by using Italian administrative data, at a very disaggregated geographical level, and suggest that, at least to some extent, the contrasting findings of the literature can be reconciled by considering the mutating information set available to the public as the pandemic progresses.

More generally, our paper contributes to the literature on the local correlates of the COVID-19 pandemic. Alacevich et al. (2021) investigate the role of care homes in Lombard municipalities hypothesising that care homes have helped spread the virus in the nearby areas. They compute excess death for each day between January 1st and March 31st and each municipality and compare municipalities with and without care homes located in their territory, controlling for other municipality characteristics and province fixed effects. Results confirm that the presence of care homes is associated with 41% larger excess death rates in the first quarter of 2020. Brandily et al. (2020) study excess mortality across municipalities in France and focus on the role of poverty. They apply a triple difference strategy, which compares the excess death in poor and non-poor municipalities located in high and low intensity areas during the first wave of the epidemic, and conclude that excess mortality was twice as large as in the poor municipalities. Also in France, Ginsburg et al. (2021) point out that departments with higher income inequality (measured by the Gini index) faced more COVID-19 deaths, more discharged patients, and a higher number of cases in the period between May 13th and September 3rd 2020. Their analysis is cross-sectional and controls for the age structure of the population, the supply of primary health care, the average household size and other local characteristics including the prevalence of COVID-19 tests. In neighbouring Belgium, Verwimp (2020) documents that COVID-19 spreads faster in municipalities that are larger, more densely populated, have higher income, are more exposed to migration, business or leisure travelling, and have a larger share of elderly people and residents in care homes. González-Val and Marcén (2021) assess the role of mass gatherings in spreading the virus in the early phases of the pandemic and estimate that one standard deviation increase in the attendance rate implied a significant increase of 3.1 COVID-19 daily cases per 100,000 inhabitants (21% of the post-effect standard deviation). The analysis is conducted at the province level and exploits the variation across provinces in the timing and attendance to mass events which were planned well in advance compared to the rise of the pandemic. Interestingly, the authors show that the degree of attendance relative to the population is only weakly, if anything, correlated with province socio-economic characteristics.

Turning to the US, Desmet and Wacziarg (2021) study the determinants of new infections and fatalities by county and conclude that population density, presence of nursing homes, lower income, higher poverty rates, and a greater presence of African Americans and Hispanics are positively correlated with the epidemic intensity and that their effects increased over time before plateauing or slightly declining. Three important contributions study New York City at the level of zip code. Glaeser et al. (2020) show that mobility is a major determinant of COVID-19 spread. Almagro et al. (2020) study the role of occupation and find the strongest positive correlation between COVID-19 prevalence and the share of workers in the transportation, industrial, natural resources and construction, and non-essential – professional sector. Borjas (2020) finds that people residing in poor or immigrant neighbourhoods were less likely to be COVID-19 tested, casting doubts on the reliability of the measures of COVID-19 prevalence.

Several papers studied the role of social capital on the pandemic. Bartscher et al. (2021) using data from seven European countries find that social capital, captured by the level of voter turnover in the European Elections of 2019, was initially associated with an increase in the number of COVID-19 cases. Next, as information on the virus spread, high-social-capital areas started to display a slower increase in COVID-19 cases. The role of social capital vanished when national lockdowns were enforced. Eventually, high-social-capital areas accumulated between 14% and 34% fewer COVID-19 cases between mid-March and late June 2020 than comparable low-social-capital areas. Likewise, high-social-capital areas recorded between 6% and

35% fewer excess deaths in Great Britain, the Netherlands, Italy, and Sweden. A similar pattern is also observed by Borgonovi et al. (2021) in the US. Durante et al. (2021) find that provinces richer of social capital complied more with mobility restrictions. For the US, Barrios et al. (2021) and Brodeur et al. (2020) confirm these results finding that social distancing was greater in areas with higher civic capital or trust.

The paper closest to ours is Armillei et al. (2021), which examines the heterogeneous COVID-19-induced excess mortality in Italian municipalities in March 2020 compared to the average March-mortality recorded between 2017 and 2019, distinguishing between central and peripheral municipalities. Their findings indicate that excess mortality was higher in peripheral than in central municipalities. However, the attributes of centrality and periphery reflect underlying differences in income, education and demographics. When the latter are included in the analysis by means of a procedure of dimension reduction, results indicate that mortality was higher in those municipalities with a lower index of income & education. Apart from focusing directly on education (rather than on the multidimensional construct of centrality, of which education is just one of the many factors), our paper differentiates from Armillei et al. (2021) for the emphasis on the time-varying effect of education in the various phases of the pandemic and the attempt to rationalize this finding.

3. Background: the epidemic in Italy

The first imported cases of COVID-19 detected in Italy dated to mid-January 2020. The first autochthone cases were reported in late February and the first fatality was dated February 21st. The virus was initially found in Lombardy and Veneto, but by early March the pandemic spread in all Northern Italy. From March 9th the whole country was locked down. Lockdown was progressively eased since May 4th to be finally lifted on June 3rd. From that moment on, travelling between regions was permitted again. During the period between June and September, mobility restrictions were dropped, economic activities re-opened and the obligation of wearing masks outdoors was removed.

From early October 2020, with a significant delay with respect to Spain and France, the first European countries that entered into the second wave, the number of infections and deaths increased once more. Restrictive measures were re-established although at a lower level of intensity compared to those enacted between March and April. Schools remained open and most firms continued to run, guaranteeing adequate preventive measures. Heavy restrictions were instead imposed on hotels, bars and restaurants and personal services. The second wave declined between January and February 2021, but it quickly left way to a third wave since March 2021. The third wave covered the period March-May 2021 and was followed in the summer by a period of low circulation of the virus, at least until September 2021. From (late) October 2021 the number of cases started to increase anew and exploded by the end of the year with the so-called Omicron variant. Our data cover the period until September 2021.

The vaccination campaign was officially launched on December 27th 2020. Its implementation was initially slow and, by the end of March 2021, only about 10% of the population received the first dose. Only by mid-June 2021, the coverage rate reached 50% (25% having received two doses). Mass vaccination represented a major game-changer, COVID fatality declined and vaccinated people earned a pass granting them full mobility and access to work, schools and public places (the non-vaccinated could earn a temporary pass after having been tested negative).

In Fig. 1, we report the excess mortality rate among the population aged over 60, by month, comparing the level of mortality rate in 2020 and 2021 with its average between 2012 and 2019.[25] We distinguish between Northern and Central-Southern Italy. It is apparent that excess mortality concentrates in two waves, between March and May in the North, and between October and December in the whole country. In the North, mortality rates were about 59% higher than usual during the first wave, and despite this experience, the North had the largest excess mortality also in the second wave (about 37% versus less than 17% in the Centre and the South). Compared to a pre-covid monthly mortality rate of about 3.2 per 1000 inhabitants aged over 60, the *excess* mortality rate in the North was over 3 in March, just below 2 in April and again around 1.5 in November and December. In the Centre and South, excess mortality rates remained below 1 between November and December.

Excess mortality was concentrated in the age group 60 + . We do not detect statistically significant variation at younger ages. In Figure C1 in the Online Appendix, we report the average excess mortality rate in Italy by age group during the first and the second wave. Among the population aged over 60, excess mortality was 1 during the first wave and 0.88 during the second. Among younger groups, excess mortality rates were very close to zero.

During the lockdown in March and April 2020, only the essential economic sectors were allowed to run. A large share of workers remained either idle or was working from home. According to Galasso and Foucault (2020), based on the data from a real-time survey carried out during the first wave, less than 40% of the low educated continued working at their usual workplace (the others were at home, in most cases idle). The proportion of those who continued to work in presence was 27% among the high school graduates and 19% among the college graduates. During the following waves, since October 2020, closures were limited to restoration and leisure services, while industry was generally spared. The less educated were more likely than the more educated to be back at the office, while home working remained widespread among the more educated.

4. Data and descriptive statistics

We match education and other socio-demographic indicators from the 2011 Census (the most recent complete Census of the population), with subsequent age-specific mortality data, recorded between 2012 and 2021, by municipality.

Mortality data are provided by ISTAT and consists of daily counts of all-cause deaths, by municipality, gender and age group, between January 2012 to September 2021. By using population yearly data by municipality, age and gender, also provided by ISTAT, we compute the mortality rate per 1000 inhabitants aged over 60, for each year, from 2012 to 2021, and each period January-February, March-May, June-September, October-December, a partition which reflects the stages of the COVID-19 pandemic in 2020 and 2021.[26] By considering mortality rates rather than death counts, we neutralize the effect of municipality scale and by focusing on the population aged 60 or more, we account for the fact that COVID-19 mortality regards almost exclusively the seniors.

In Fig. 2.a (resp. 2.b), we display mortality rates in the population aged 60 and over, by quintile of municipal education in Northern (resp. Centre-Southern) Italy, separately for each of the four periods and for all years between 2012 and 2021. Our main measure of education is the share of municipal residents with at least secondary education in 2011, predetermined with respect to all mortality data.[27]

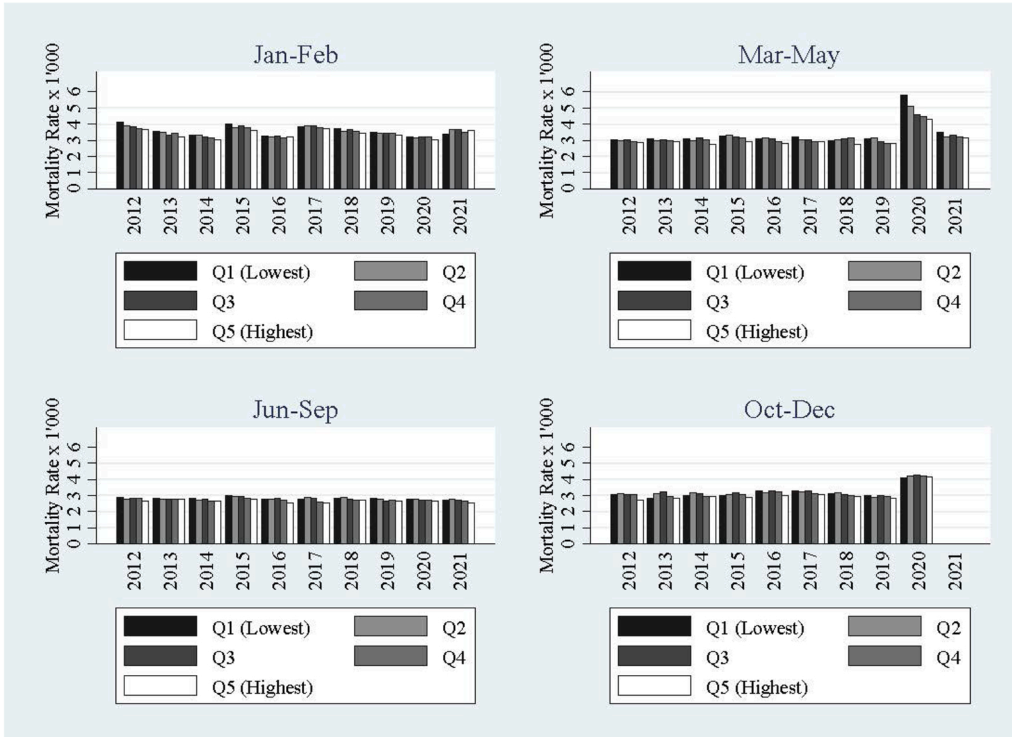
In the North, the upsurge of mortality in the periods March-May 2020 (first COVID-19 wave) and October-December 2020 and January-February 2021 (second COVID-19 wave) is evident when

[25] Such situation was not specific to Italy but occurred in many other countries (Zhang et al., 2021).

[26] President, Consiglio Superiore della Sanità.

[27] Many studies use the word infodemic to describe the wave of unscientific claims that accompanied the pandemic (see Gallotti et al. 2020 among others). A list of “popular" fake news has been collected by the Ministry of Health and confuted at the following address: https://www.salute.gov.it/portale/nuovocoronavirus/archivioFakeNewsNuovoCoronavirus.jsp

a



b

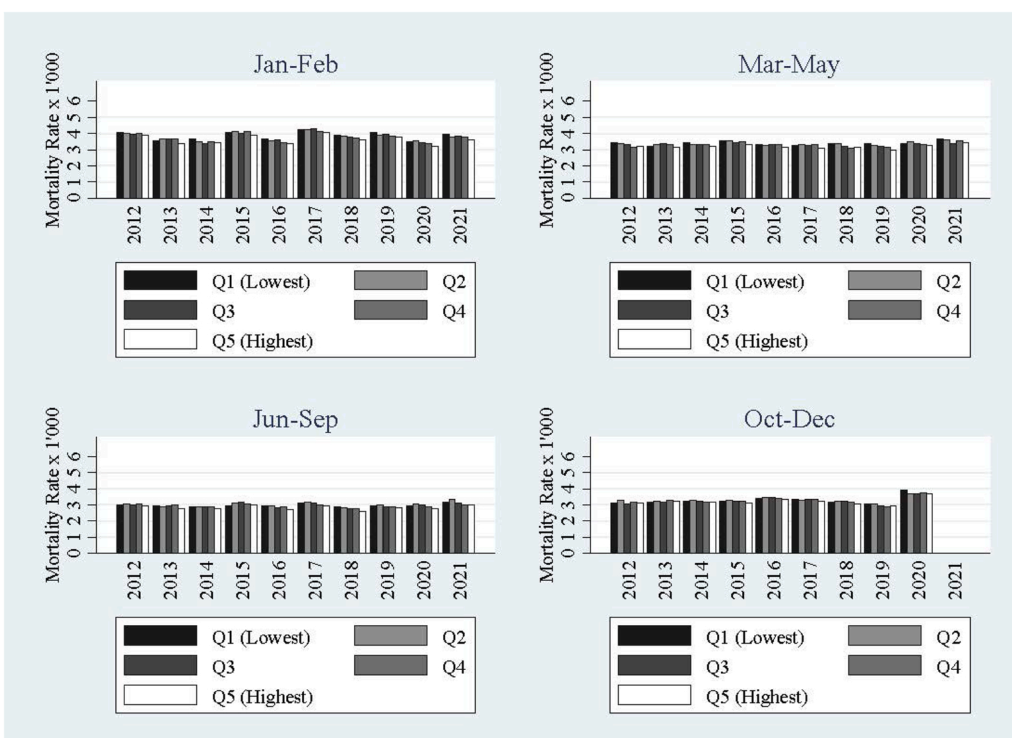
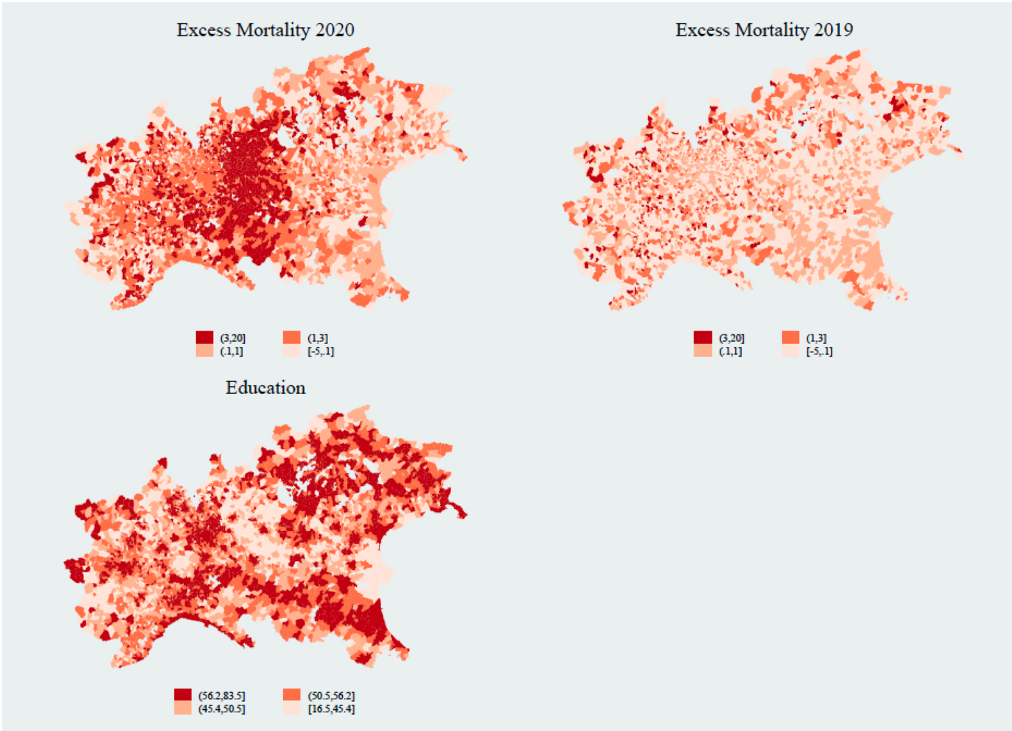


Fig. 2. a: Average Mortality Rate, by education (2012–2021) – North. b: Average Mortality Rate, by education (2012–2021) – Centre- South. *Note: The figures show the evolution of municipality average mortality rate between 2012 and 2021, separately for Northern and Central-Southern Italy and period. The figure also distinguishes by quintile of municipal education* (i.e. *share of residents with at least upper secondary school). Source: ISTAT.*

a



b

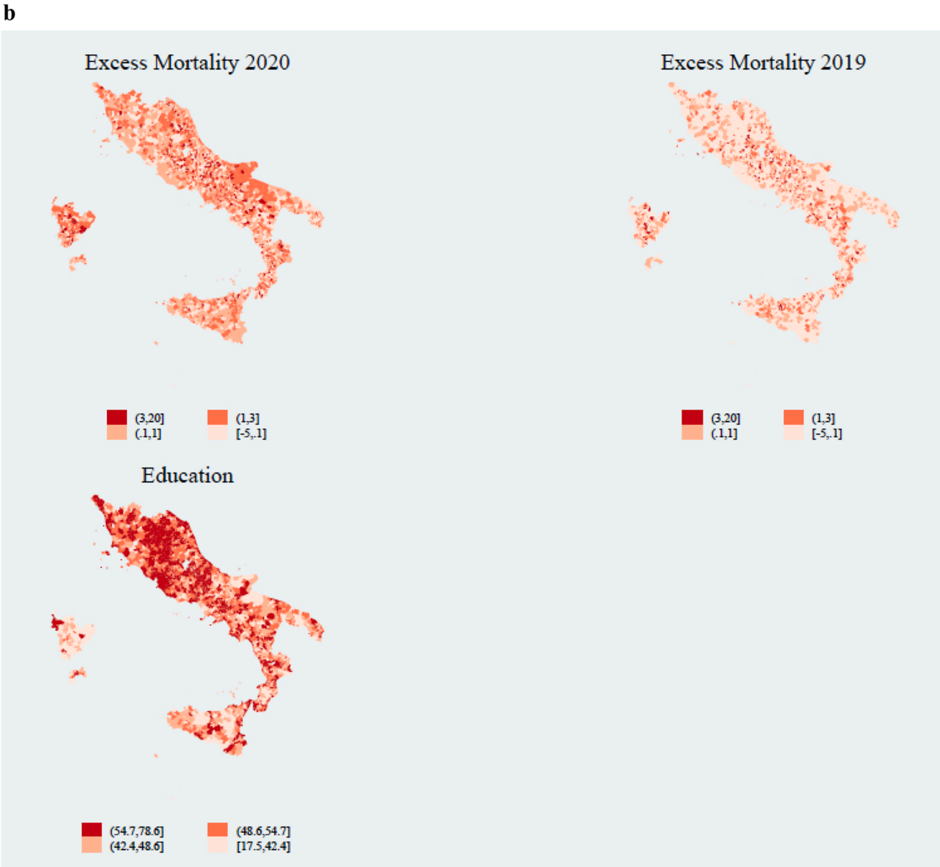


Fig. 3. a: **Excess Mortality and Education, North, Mar-May 2020 and 2019.** Note: The figures show the distribution of excess mortality and education across municipalities in North Italy during the first wave of the pandemic and in the same period in 2019. Education is measured by the share of individuals with upper secondary education. **b: Excess Mortality and Education, Centre-South, Oct-Dec 2020 and 2019.** *Note: The figures show the distribution of excess mortality and education across municipalities in Centre-South Italy during the second wave of the pandemic and in the same period in 2019. Education is measured by the share of individuals with upper secondary education.*

Table 1
Summary Statistics.

	# municipalities	mean	std dev.	min	max
Monthly Mortality Rate (2021) x 1000 inhabitants	7550	3.316	1.201	0.000	14.403
Monthly Mortality Rate (2020) x 1000 inhabitants	7550	3.571	1.210	0.000	16.667
Monthly Mortality Rate (2012–19) x 1000 inhabitants	7550	3.190	0.673	1.433	9.040
Share of residents with at least high school	7550	49.703	8.831	16.500	83.500
Share of college graduates	7550	8.307	3.324	0.401	33.748
Employment rate	7550	45.157	7.905	18.000	74.000
Index of commuting	7550	0.802	0.071	0.220	0.960
Number of hospital beds per 1000 inhabitants	7550	1.095	10.433	0.000	684.100
Incidence of house ownership	7550	76.662	6.609	17.600	100.000
Housing price index	7550	0.783	0.438	0.172	7.509
Share of population out of the main center	7550	17.850	18.274	0.000	97.400
Share of families at risk of poverty	7550	2.028	1.884	0.000	17.900
Share of migrant population	7550	5.893	4.192	0.000	36.700
Population density	7550	287.149	616.184	1.400	11346.300
Average family size	7550	2.360	0.265	1.200	3.400
Ratio of population older than 64 to population younger than 15	7550	194.232	138.542	25.400	2850.000
Male/female ratio	7550	97.044	6.242	67.800	182.800
Average population age	7550	44.424	3.998	21.276	68.768
Resident population	7550	7658	41096	30	2617175
Average temperature 2009–2018	7550	13.384	2.866	-0.193	20.503
Hotel beds per 1000 inhabitants	7550	0.165	0.542	0.000	10.572
Number of places in care homes per 1000 inhabitations	7550	16.377	12.267	0.187	43.058
Voter turnout in European Elections 2009	7550	0.697	0.139	0.068	1.061
Distance from the nearest airport (meters)	7550	42204.809	23530.507	269.011	143222.594
Degree of teleworkability	7550	0.338	0.045	0.239	0.468

Note: The table shows summary statistics for our dependent and control variables. All variables are at the municipal level, except for the number of nursing home beds which is at the regional level. An observation is a municipality-period-year. Source: ISTAT, ELIGENDO (Archivio Storico delle Elezioni) and UERRA.

compared to the previous years. Excess mortality between March and May 2021 (third wave) although positive is substantially smaller than in the earlier waves. Partly this result is to be ascribed to the vaccination campaign, which initially targeted the most seniors (the over 80 were the first to be vaccinated, followed by the over 70) and reached relatively high coverage rates already by May 2021. An education gradient is marked during the first wave as the municipalities in the lowest quintile of the education distribution reached a mortality level of almost 6 per 1000 inhabitants compared to a mortality rate slightly above 4 in the municipalities in the highest quantile (see Table C1 in the Online Appendix). During the second and third waves, differences by education almost disappeared as mortality rates are around 4 in all quantiles. In the Centre-South, a moderate excess mortality level is recorded only between October and December 2020 with only a hint of education gradient. In the following wave, excess mortality is rather small and unrelated to education.

Both figures show that in pre-COVID years mortality is higher, on average, and more volatile from year to year, in the January-February period than in other periods of the year. This is the so-called “excess winter mortality” documented among others in Healy (2003) and Lerchl (1998). Mortality rates range between 3 and slightly above 4 per 1000 inhabitants, while they remain around 3 per 1000 inhabitants in the rest of the year. Especially in the North, in January-February, we observe an education gradient that is instead absent or much more nuanced in the other three periods. Higher mortality and higher volatility are the result of, among others, the seasonal flu, whose strength varies from year to year (Rolfes et al., 2018) and the variation in winter low temperatures, which are associated with a number of health disorders (Lerchl, 1998). The education gradient in the seasonal flu has been previously observed in Crighton et al. (2007) and an association between education and the propensity to vaccinate against influenza is documented, for instance, in Mills et al. (2016) and Nagata et al. (2013).

Moreover, by focusing on North-Italy and on the first wave of the pandemic, Figs. 3a and 3b provides two maps showing the distribution of excess mortality and education across municipalities in the North and Centre-South municipalities, respectively. A visual inspection of these maps confirms the evidence discussed above that municipalities with a higher average level of education have experienced lower excess mortality.[28]

In the empirical analysis below, the outcome variable will be excess mortality by municipality, period and year, among the population aged 60 and over. Excess mortality is the difference between mortality rates recorded in 2020 (resp. 2021) with respect to the average mortality rates between 2012 and 2019 (by period and municipality).[7] Excess mortality has a number of advantages over other possible indicators. First, it is to be preferred with respect to COVID-19 mortality because it is not affected by miss-classifications of the cause of death, and because there might be indirect fatalities related to the pandemic. The latter include deaths due to congestion of the healthcare services, but also possibly averted deaths due to a smaller number of road traffic accidents or work accidents. Similarly, it is to be preferred over any indicator depending on the number of COVID-19 cases, because COVID-19 cases are largely underestimated (Wu et al., 2020; Irons and Raftery, 2021).[8] Moreover excess mortality is typically available at a finer geographical level, as it is based on data from municipality population registries. Other possible interesting outcomes, such as the number of COVID-19 hospitalizations or the number of hospitalizations in intensive care, are originally recorded at the hospital level and then aggregated by province or region.[9] Finally, excess mortality combines the spread of the virus, its

[28] Twitter does not allow to distinguish by area of origin. We isolate tweets written in Italian and we used the Italian translation of “Vitamine C” and “Bleach”.

[7] As a robustness we compared 2020 (or 2021) with the average between 2015 and 2019 – see Section 6.

[8] COVID-19 remains often undetected among the asymptomatic, or because infected people prefer not to not report to their authorities their status.

[9] Also for hospitalization, there is a moderate risk of miss-classification, as COVID-19 positives can be hospitalized for reasons different than COVID (this occurs in about 10% of the cases – Tsai et al., 2021), and can be classified as hospitalization for COVID-19 (incident COVID hospitalizations).

fatality, the adoption and effectiveness of preventive and care strategies. In this sense, it is a comprehensive measure of the epidemic burden.

Municipality socio-demographic indicators have been compiled by the Local Opportunities Lab,[10] a think tank, which harmonizes data by municipality produced by ISTAT and other government agencies. We select census data for 2011, which include the share of individuals with at least upper secondary education, employment rate, an index of commuting, the share of families at risk of poverty, the share of migrant population, population density, the incidence of house ownership, a housing price index, the number of hospital beds per inhabitant, the number of hotel beds per inhabitant, the share of population living out of the main agglomeration of the municipality, average family size, the dependency ratio (i.e. the ratio of the population older than 65 to the population younger than 15), average population age, the male to female ratio, and the same indicator of social capital recently adopted by Bartscher et al. (2021), the voter turnout at the European Elections of 2009.[11] Moreover we collect the 2008–2019 municipal average temperature from UERRA,[12] and the planar distance between each municipality and the closest airport. Finally, at a higher level of geographical aggregation, we collect data on the number of nursing home beds per inhabitant by region, and construct an index of teleworkability by province using the Italian LFS.[13] Table 1 provides summary statistics for all these variables. After dropping the municipalities with missing data, the remaining sample includes 96% of all municipalities (7550) and 97% of Italian population.

5. Empirical analysis and results

We hypothesise that the causal effect of COVID-19 on mortality is possibly modified (moderated) by education. The goal of this and the following sections is that of testing whether and to what extent this is the case, while Section 7 discusses the rationale and the possible mechanisms underlying the role played by education.

We run separate regressions for Northern and Central-Southern Italy, and estimate the effect of education on excess mortality in each of the periods January-February, March-May and June-September of both 2020 and 2021 and October-December 2020, to identify how the role of education changed across the phases of the pandemic. Indeed, March-May 2020 corresponds to the first wave, October-December 2020 and January-February 2021 to the second wave, and, March-May 2021 to the third wave. All others are periods of low epidemic intensity.

Conducting analyses within period, rather than exploiting variation between periods, has two advantages. First, given the seasonal differences in the education gradient of mortality documented in Figs. 2a and 2b and related to the "excess winter mortality", the pre-pandemic months of January and February 2020 would not be a good counterfactual for the post-pandemic months of March to May, nor for any other period of the year. Second, by comparing the total number of deaths of a wave with that of the corresponding period in previous years, we address the concern raised by Callaway and Li (2021) and Bisin and Moro (2020) that comparing mortality within and across waves might yield biased estimates, because of strong nonlinearities in the evolution of the pandemic. Moreover, by conducting the analysis by area, we address the concern of Callaway and Li (2021) that different areas can be at different stages of the epidemic at a given date. Indeed, as we already remarked, the Centre-South was spared from the first wave and in October-December 2020 it experienced the first major outbreak, while the North was already at the second.

We define

$$\Delta M_{itp} = \ddot{\alpha}_{pt} + \ddot{\beta}_{pt} Edu_i + \ddot{\delta}_{pt} X_i + \ddot{\varepsilon}_{ipt} \tag{1}$$

where $\Delta M_{itp} = M_{itp} - \overline{M}_{ip}^{12-19}$ is the difference between the mortality rate in period $p \in \{Jan-Feb, Mar-May, Jun-Sept, Oct-Dec\}$ in year $t \in \{2020, 2021\}$ and the average pre-COVID mortality rate between 2012 and 2019 in the same period of the year for municipality i (excess mortality);[14] Edu_i is the municipal share of residents with at least upper secondary education in 2011 and X_i is the set of predetermined municipal control variables; $\ddot{\varepsilon}_{ipt}$ is the residual, allowed to be clustered within local labour systems (a concept akin to that of commuting area) to account for heteroskedasticity and spatial correlation.

Model (1) is estimated by a fully interacted regression with Edu_i and X_i interacted by a full set of period-year dummies, so that parameter identification does not involve any comparison across period-years. For each period-year the parameters $\ddot{\beta}_{pt}$ identify the effect of education on COVID-19-induced excess mortality, provided that the *trends* of mortality levels, absent the pandemic, are parallel between municipalities with different education levels. Otherwise, excess mortality would be associated to education also for reasons other than the role of education in the COVID-19 epidemic.

We refer the reader to Appendix A for more details about the parallel trend assumption and the derivation of model (1). As reported in Table A1, we test and never reject the hypothesis of parallel trend by using mortality data from the period between 2012 and 2019. This result is not surprising given that mortality in the row data depicted in Figs. 2a and 2b appears flat between 2012 and 2019. Indeed, not only mortality trends are parallel between municipalities with different levels of education, but mortality levels are practically constant over time.

Model (1) describes how the effect of COVID-19 on mortality varies depending on predetermined municipality education and the choice of controls is motivated by the need of neutralizing possible confounders, rather than providing a complete account of all the possible dimensions of heterogeneity. The list of controls is included in Table 1. We account for 1) urbanization level (captured by population density and population size), which likely puts municipalities on divergent mortality trends, because of differential air pollution and lifestyles, and it is shown to influence the spread of the COVID-19 epidemic (Armillei et al., 2021; Desmet and Wacziarg, 2021); 2) the level of commuting from nearby areas (captured by the share of population out of the main center, and an index of commuting), which also contributes to virus spread (Glaeser et al., 2020); 3) the level of employment and the type of occupation (captured by the employment rate, and the incidence of teleworkable occupations) which accounts for the differential bite of mobility restrictions (Almagro et al., 2020); 4) affluence and income distribution

[10] https://www.localopportunitieslab.it/

[11] We consider elections of 2009 to be consistent with the timing of other controls. Bartscher et al. (2021) consider voter turnout at the provincial level, while our measure is much finer, being defined by municipality. Alternatively, we used three variables that are commonly adopted in the literature, the incidence of blood donation, the turnout in the referendum on divorce in 1974 (Guiso, Sapienza, and Zingales, 2004) and the answer to 'trust' question in the World Value Survey (Tabellini, 2010). All variables are by province. Results, available upon request, do not chance qualitatively.

[12] The data on the municipal average temperature comes from the UERA dataset, which includes surface and near-surface essential climate variables from UERRA-HARMONIE and MESCAN-SURFEX systems, for almost all European countries.

[13] The teleworkability index comes from Sostero et al. (2020), who identify the jobs that can be done at home and those that cannot and provide a measure of teleworkability by ISCO-08 code. We use the 2019 Italian Labour Force survey, which includes the information on the occupation and province of residence of the individual, to calculate the average index at the province level in Italy.

[14] An advantage of this specification is that excess mortality is measured with respect to the average mortality between 2012 and 2019, which is less affected by random noise compared to mortality in 2019. A similar specification is adopted for instance in Bartscher et al. (2021). However even by taking 2015–2019 as the baseline period against which comparing 2020 and 2021 mortality does not appreciably change our results.

Table 2
The effect of education on excess mortality. By stage of the COVID-19 pandemic. North.

	(1)	(2)	(3)	(4)	(5)	(6) Average excess mortality rate [average mortality rate]
Edu * Jan-Feb 20	0.00114	0.00359	0.00270	0.00354	0.00385	-0.396
	(0.00623)	(0.00694)	(0.00695)	(0.00842)	(0.00845)	[3.140]
Edu * Mar-May 20	-0.0482 * **	-0.0577 * **	-0.0569 * **	-0.0479 * **	-0.0426 * **	1.783
	(0.0146)	(0.0172)	(0.0163)	(0.0147)	(0.0139)	[4.789]
Edu * Jun-Sep 20	0.00122	-0.00250	-0.00224	-0.00146	-0.000830	-0.041
	(0.00413)	(0.00516)	(0.00518)	(0.00664)	(0.00638)	[2.717]
Edu * Oct-Dec 20	0.00923	0.00255	0.00222	0.0100	0.00764	1.153
	(0.00632)	(0.00686)	(0.00677)	(0.00773)	(0.00757)	[4.197]
Edu * Jan-Feb 21	0.0165 * **	0.0117 *	0.0103	0.00839	0.00688	0.030
	(0.00562)	(0.00688)	(0.00675)	(0.00829)	(0.00838)	[3.565]
Edu * Mar-May 21	-0.00181	-0.00674	-0.00668	-0.00891	-0.00867	0.250
	(0.00493)	(0.00550)	(0.00547)	(0.00704)	(0.00711)	[3.255]
Edu * Jun-Sep 21	0.000295	-0.000472	-0.000156	-0.00171	-0.000773	0.085
	(0.00416)	(0.00506)	(0.00503)	(0.00584)	(0.00594)	[2.67]
Observations	29,718	29,718	29,718	29,718	29,718	
R-squared	0.113	0.123	0.132	0.140	0.144	
Controls I*Period-Year Dummies	NO	YES	YES	YES	YES	
Controls II*Period-Year Dummies	NO	NO	YES	YES	YES	
Controls III*Period-Year Dummies	NO	NO	NO	YES	YES	
Controls IV*Period-Year Dummies	NO	NO	NO	NO	YES	

Note: Years: 2012–2021. Robust standard errors clustered at the local labour system in parentheses. The dependent variable is excess mortality, defined as the difference between the monthly mortality rate for the age group aged 60 + in the period/year reported in rows and the corresponding average for the same period between 2012 and 2019. Edu measures the share of individuals with at least upper secondary education. All period-year models are stacked jointly estimated by means of a fully interacted regression, whose R2 is reported in the table. Controls I: average population age, population density, family size, share of migrants, male to female ratio, share of population older than 64; Controls II: mobility, share of population out of main center, distance from the airport; Controls III: hospital beds, nursing homes, share of families in poverty, hotel beds, household ownership, house price index, employment, teleworkability index, average temperature; Controls IV: Voter turnout. * ** $p < 0.01$, * * $p < 0.05$, * $p < 0.1$.

Table 3
The effect of education on excess mortality. By stage of the COVID-19 pandemic. Centre-South.

	(1)	(2)	(3)	(4)	(5)	(6) Average excess mortality rate [average mortality rate]
Edu * Jan-Feb 20	-0.00247	-0.00280	-0.00355	-0.00389	-0.00513	-0.389
	(0.00478)	(0.00532)	(0.00552)	(0.00658)	(0.00727)	[3.443]
Edu * Mar-May 20	-0.00000	-0.000992	-0.00139	0.000605	0.000940	0.096
	(0.00359)	(0.00407)	(0.00411)	(0.00452)	(0.00460)	[3.396]
Edu * Jun-Sep 20	-0.00325	-0.00118	-0.00102	0.000251	0.000997	0.004
	(0.00265)	(0.00297)	(0.00309)	(0.00370)	(0.00372)	[2.947]
Edu * Oct-Dec 20	-0.00741 * *	-0.00901 * *	-0.00879 * *	-0.00867 * *	-0.00811 * *	0.561
	(0.00354)	(0.00360)	(0.00361)	(0.00384)	(0.00390)	[3.773]
Edu * Jan-Feb 21	-0.00656	-0.00715	-0.00848	-0.00479	-0.00453	-0.017
	(0.00552)	(0.00556)	(0.00554)	(0.00629)	(0.00633)	[3.815]
Edu * Mar-May 21	-0.00619 *	-0.00439	-0.00414	-0.00442	-0.00471	0.273
	(0.00367)	(0.00411)	(0.00415)	(0.00471)	(0.00474)	[3.573]
Edu * Jun-Sep 21	-0.00654 * *	-0.00654 * *	-0.00556 *	-0.00343	-0.00252	0.199
	(0.00263)	(0.00316)	(0.00317)	(0.00373)	(0.00384)	[3.143]
Observations	23,121	23,121	23,121	23,121	23,121	
R-squared	0.026	0.031	0.032	0.037	0.038	
Controls I*Period-Year Dummies	NO	YES	YES	YES	YES	
Controls II*Period-Year Dummies	NO	NO	YES	YES	YES	
Controls III*Period-Year Dummies	NO	NO	NO	YES	YES	
Controls IV*Period-Year Dummies	NO	NO	NO	NO	YES	

Note: Years: 2012–2021. Robust standard errors clustered at the local labour system in parentheses. The dependent variable is excess mortality, defined as the difference between the monthly mortality rate for the age group aged 60 + in the period/year reported in rows and the corresponding average for the same period between 2012 and 2019. Edu measures the share of individuals with at least upper secondary education. All period-year models are stacked jointly estimated by means of a fully interacted regression, whose R2 is reported in the table. Controls I: average population age, population density, family size, share of migrants, male to female ratio, share of population older than 64; Controls II: mobility, share of population out of main center, distance from the airport; Controls III: hospital beds, nursing homes, share of families in poverty, hotel beds, household ownership, house price index, employment, teleworkability index, average temperature; Controls IV: Voter turnout. * ** $p < 0.01$, * * $p < 0.05$, * $p < 0.1$.

(captured by the incidence of house ownership, the housing price index, and share of families at risk of poverty) which influence individual ability to self-protect (Brandily et al., 2020; Ginsburgh et al., 2021); 5) demographics (captured by average age, dependency ratio, male/female ratio, average family size, and share of migrant population), as the case-fatality ratio is higher among the seniors and intergenerational residence patterns influence virus spread within families (Aparicio Fenoll and Grossbard, 2020; Verwimp, 2020); 6) healthcare facilities (captured by number of hospital beds, number of places in nursing homes), as institutions can turn into hotspots for the virus (Alacevich, 2021); 7) travelling facilities (captured by the distance from the nearest airport and the number of bed places in hotels) as long-range mobility determines how fast the virus reach an area (Daon et al., 2020; Notari and Torrieri, 2021); 8) the level social capital, (captured by voter

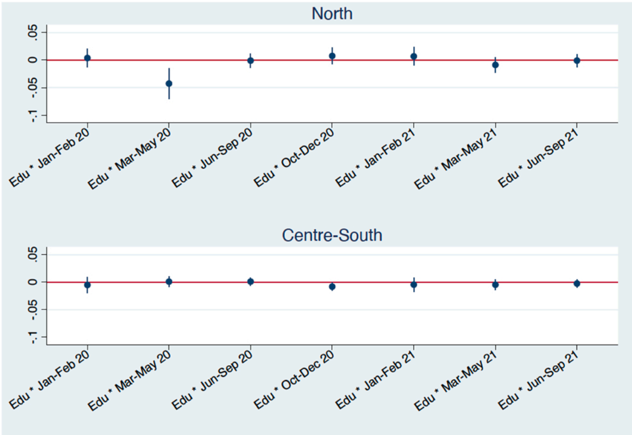


Fig. 4. Estimates of the differential effect of education between Jan-Feb 20 and Jun-Sep 21 (model 1). Note: The figures show the coefficients of the interaction terms between Education and the stage dummies from Model 1 by stage of the pandemic.
Source: ISTAT.

turnout by municipality in the 2009 European Elections) as social capital influences influence the intensity of social relationships and compliance with public health regulations (Bartscher et al., 2021; Durante et al., 2021; Borgonovi et al., 2021); and 9) climatic conditions (captured by the average temperature between 2009 and 2018), which is associated with virus ability to spread out (Mecenas et al., 2020).

6. Results

Estimates of Model (1) by area and by period-year, are reported in Table 2 (North) and Table 3 (Centre-South). In each Table, we start from the most parsimonious specification in column (1) where controls are omitted, and then we progressively include controls in columns (2) to (5). Finally, column (6) reports mortality levels and excess mortality to help gauge the magnitude of education effects. Fig. 4 provides a graphical representation of the findings corresponding to the richest specification of the model. For Northern Italy, we find evidence that education significantly reduced excess mortality in the period March-May 2020, but not in any other period. Between March and May 2020, the average excess mortality rate in Northern Italy was 1.783 deaths per month, compared to an average mortality rate of 3.140 out of 1000 inhabitants aged 60 + in the pre-COVID times. We find that an increase of 10% points in the share of residents with at least secondary education corresponds to a reduction of excess mortality of 0.426 deaths per month, equivalent to 24% of the average excess mortality.[15] In Centre-Southern Italy, we find a smaller negative effect in October-December 2020. In this area and period-year, excess mortality was 0.561 compared to a pre-COVID mortality rate of 3.443 and an increase of 10% points in the share of residents with at least secondary education corresponds to a reduction in excess mortality of 0.081 deaths per month, equivalent to 14% of the average excess mortality.

Results are little affected by the inclusion of municipal controls. Indeed, controls turn out to be generally insignificant and contribute little to the proportion of explained variation. A notable exception is voter turnover in the European elections of 2019. Our analysis confirms the pattern found by Bartscher et al. (2021), whereby social capital initially favours the diffusion of the epidemic while later on it helps containing it. It is worth noting that our analysis is at the municipal level while Bartscher et al.'s is by province. Detailed results for the role of social capital are reported in the Online appendix Table C2. Also in this case the effects of social capital are more pronounced in the North than in the Centre-South. In the former area, in the period March-May 2020, the effect of social capital is opposite to that of education and their magnitude is comparable in absolute value. Hence while education was protective, social capital favoured the spread of the epidemic. In the

Table 4
The effect of education on excess mortality. By Area. Largest and smallest municipalities excluded.

VARIABLES	Smallest/ largest 1% excluded North	Smallest/ largest 1% excluded Centre-South	Smallest/ largest 5% excluded North	Smallest/ largest 5% excluded Centre-South
Edu * Jan-Feb 20	0.00333	-0.00365	-0.00172	-0.00662
	(0.00901)	(0.00666)	(0.00689)	(0.00614)
Edu * Mar-May 20	-0.0426 * **	8.18e-05	-0.0531 * **	0.000629
	(0.0144)	(0.00459)	(0.0142)	(0.00480)
Edu * Jun-Sep 20	0.000435	-0.000263	-0.00139	-0.00359
	(0.00465)	(0.00367)	(0.00437)	(0.00390)
Edu * Oct-Dec 20	0.00227	-0.00638	0.00489	-0.00756 *
	(0.00814)	(0.00406)	(0.00796)	(0.00404)
Edu * Jan-Feb 21	0.00837	-0.00463	0.000610	-0.00344
	(0.00827)	(0.00596)	(0.00672)	(0.00569)
Edu * Mar-May 21	-0.00838	-0.00355	0.00190	-0.00154
	(0.00672)	(0.00448)	(0.00568)	(0.00468)
Edu * Jun-Sep 21	-0.00289	-0.00344	-0.00994 * *	-0.00332
	(0.00529)	(0.00369)	(0.00479)	(0.00363)
Observations	29,017	22,750	26,684	20,867
R-squared	0.153	0.037	0.186	0.039
Controls I*Period-Year Dummies	YES	YES	YES	YES
Controls II*Period-Year Dummies	YES	YES	YES	YES
Controls III*Period-Year Dummies	YES	YES	YES	YES
Controls IV*Period-Year Dummies	YES	YES	YES	YES

Note: Years: 2012–2021. In columns 1 (3) and 2 (4), the largest and the smallest 1 (5) percent of the municipality distribution is excluded from the sample. Robust standard errors clustered at the local labour system level in parentheses. The dependent variable is excess mortality, defined as the difference between the monthly mortality rate for the age group aged 60 + in the period/year reported in rows and the corresponding average for the same period between 2012 and 2019. Edu measures the share of individuals with at least upper secondary education. All period-year models are stacked jointly estimated by means of a fully interacted regression, whose R2 is reported in the table. Controls I: average population age, population density, family size, share of migrants, male to female ratio, share of population older than 64; Controls II: mobility, share of population out of main center, distance from the airport; Controls III: hospital beds, nursing homes, share of families in poverty, hotel beds, household ownership, house price index, employment, teleworkability index, average temperature; Controls IV: Voter turnout.* ** $p < 0.01$, * * $p < 0.05$, * $p < 0.1$.

[15] Estimates are similar if the time span of the first wave is shortened to include only March and April.

Table 5
The effect of education on excess mortality. By Area. Pre-COVID period 2015–2019.

VARIABLES	(1) North	(2) Centre-South
Edu * Jan-Feb 20	0.00189	-0.00476
	(0.00842)	(0.00808)
Edu * Mar-May 20	-0.0407 * **	0.00325
	(0.0141)	(0.00469)
Edu * Jun-Sep 20	0.00103	0.00214
	(0.00650)	(0.00372)
Edu * Oct-Dec 20	0.00610	-0.00518
	(0.00771)	(0.00417)
Edu * Jan-Feb 21	0.00481	-0.00416
	(0.00855)	(0.00698)
Edu * Mar-May 21	-0.00681	-0.00240
	(0.00721)	(0.00476)
Edu * Jun-Sep 21	0.00109	-0.00137
	(0.00601)	(0.00386)
Observations	29,718	23,121
R-squared	0.135	0.039
Controls I*Period-Year Dummies	YES	YES
Controls II*Period-Year Dummies	YES	YES
Controls III*Period-Year Dummies	YES	YES
Controls IV*Period-Year Dummies	YES	YES

Note: Years: 2015–2021. Robust standard errors clustered at the local labour system level in parentheses. The dependent variable is excess mortality, defined as the difference between the monthly mortality rate for the age group aged 60 + in the period/year reported in rows and the corresponding average for the same period between 2015 and 2019. Edu measures the share of individuals with at least upper secondary education. All period-year models are stacked jointly estimated by means of a fully interacted regression, whose R2 is reported in the table. Controls I: average population age, population density, family size, share of migrants, male to female ratio, share of population older than 64; Controls II: mobility, share of population out of main center, distance from the airport; Controls III: hospital beds, nursing homes, share of families in poverty, hotel beds, household ownership, house price index, employment, teleworkability index, average temperature; Controls IV: Voter turnout. * ** $p < 0.01$, * * $p < 0.05$, * $p < 0.1$.

Table 6
The effect of education on excess mortality. By Education Quantile and By Area.

	(1) North	(2) Centre-South
Edu Q1 * Mar-May 2020	1.120 * **	
	(0.344)	
Edu Q2 * Mar-May 2020	0.529 * *	
	(0.265)	
Edu Q3 * Mar-May 2020	0.143	
	(0.193)	
Edu Q4 * Mar-May 2020	0.112	
	(0.167)	
Edu Q1 * Oct-Dec 2020		0.227 * *
		(0.102)
Edu Q2 * Oct-Dec 2020		-0.0725
		(0.0907)
Edu Q3 * Oct-Dec 2020		0.0211
		(0.0911)
Edu Q4 * Oct-Dec 2020		0.0551
		(0.0887)
Observations	29,718	23,121
R-squared	0.138	0.034
Controls I*Period-Year Dummies	YES	YES
Controls II*Period-Year Dummies	YES	YES
Controls III*Period-Year Dummies	YES	YES
Controls IV*Period-Year Dummies	YES	YES

Note: Years: 2012–2021. Robust standard errors clustered at the local labour system in parentheses. The dependent variable is excess mortality, defined as the difference between the monthly mortality rate for the age group aged 60 + in the period/year reported in rows and the corresponding average for the same period between 2012 and 2019. Edu Qx are the quintiles of the share of individuals with at least upper secondary education. The reference education quintile is the fifth (Q5). All period-year models are stacked jointly estimated by means of a fully interacted regression, whose R2 is reported in the table. Controls I: average population age, population density, family size, share of migrants, male to female ratio, share of population older than 64; Controls II: mobility, share of population out of main center, distance from the airport; Controls III: hospital beds, nursing homes, share of families in poverty, hotel beds, household ownership, house price index, employment, teleworkability index, average temperature; Controls IV: Voter turnout. * ** $p < 0.01$, * * $p < 0.05$, * $p < 0.1$.

Centre-South the effect of social capital is always small and insignificant.[16]

Unsurprisingly, there are no differential effects among municipalities in the period January-February 2020, before the arrival of the virus, and in the summer period between June and September 2020 and 2021, when the virus circulation was largely reduced. Nor we do detect effects in the North during the second wave of the epidemic between October 2020 and February 2021 and during the third wave, between March and May 2021. In this case, excess mortality is not correlated with education, the point estimate of Edu_i being very small.

6.1. Robustness checks and extensions

In this subsection, we perform some robustness checks. First, we trim the sample by dropping the largest and smallest 1% (resp. 5%) of the municipalities, in terms of their population. The smallest municipalities are likely to be rural or mountain, relatively more isolated and hence outside the main avenues of virus diffusion, while the large urban areas might have experienced harsher conditions during the epidemic and the weeks of lockdown, because of a higher population density. Results reported in Table 4 are qualitatively similar to the baseline, although the effect of education in Centre-South Italy for October-December 2020 is not statistically significant at the conventional levels in column 2.

Second, we redefine excess mortality as the difference between 2020 (2021) mortality rate and the average mortality rate between 2015 and 2019, to have a more accurate indication of the baseline mortality pre-pandemic, albeit possibly noisier. Also in this case, results reported in Table 5 do not change much compared to the baseline and also in this case the effect in the Centre-South in October-December 2020 is not precisely estimated.

Third, we explore possible non-linear effects of education. We replaced *Edu* with a set of dummies corresponding to the quintiles of the education distribution (the highest quintile is the reference category and is omitted). Results are reported in Table 6 for the periods March-May 2020 in the North and the period October-December 2020 in the Centre-South. For the North, there is a monotonic decrease in the effect of education between the first and the last quintile, with evidence of non-linearities, especially between the first and the second quantile. Moving from the first to the second quantile corresponds to a decline of mortality of 0.591 deaths per 1000 inhabitants aged 60 +, which is comparable to the decline between the second and the fifth quantile. In the Centre-South, there is practically no effect of education between the second and the fifth quantile. Excess mortality is concentrated in the municipalities belonging to the lowest quintile of the distribution.[17]

Fourth, we consider another measure of education, more restrictive, namely the share of college graduates. Overall, the effects of education turn to be stronger (Table 7). In this case, In the North, an increase of 10% points in the share of college graduates would correspond to about 0.566 lower excess mortality, per month, between March and May 2020,

[16] Also for voter turnout there is clear evidence of a common trend in the pre-pandemic period.

[17] This effect is robust the exclusion of the smallest and the largest municipalities and to the alternative definition of excess mortality. The results are available upon request.

Table 7
The effect of the share of college graduates on excess mortality. By Area.

	(1) North	(2) Average excess mortality rate [avg. mortality rate]	(3) Centre-South	(4) Average excess mortality rate [avg. mortality rate]
Col * Jan-Feb 2020	0.000924	-0.396	0.00462	-0.389
	(0.0170)	[3.140]	(0.0197)	[3.443]
Col * Mar-May 2020	-0.0566 * *	1.783	0.00246	0.096
	(0.0247)	[4.789]	(0.0115)	[3.396]
Col * Jun-Sep 2020	0.00810	-0.041	0.0136	0.003
	(0.0108)	[2.717]	(0.00929)	[2.947]
Col * Oct-Dec 2020	-0.00143	1.153	-0.0188	0.561
	(0.0174)	[4.197]	(0.0133)	[3.773]
Col * Jan-Feb 2021	-0.0132	0.030	-0.00676	-0.017
	(0.0194)	[3.565]	(0.0163)	[3.815]
Col * Mar-May 2021	-0.00772	0.250	-0.00619	0.273
	(0.0147)	[3.255]	(0.0128)	[3.573]
Col * Jun-Sep 2021	-0.0105	0.085	-0.000991	0.199
	(0.0112)	[2.67]	(0.0110)	[3.143]
Observations	29,718		23,121	
R-squared	0.143		0.037	
Controls I *Period-Year Dummies	YES		YES	
Controls II *Period-Year Dummies	YES		YES	
Controls III *Period-Year Dummies	YES		YES	
Controls IV *Period-Year Dummies	YES		YES	

Note: Years: 2012–2021. Robust standard errors clustered at the local labour system in parentheses. The dependent variable is excess mortality, defined as the difference between the monthly mortality rate for the age group aged 60 + in the period/year reported rows and the corresponding average for the same period between 2012 and 2019. Col measures the share of college graduates. All period-year models are stacked jointly estimated by means of a fully interacted regression, whose R2 is reported in the table. Controls I: average population age, population density, family size, share of migrants, male to female ratio, share of population older than 64; Controls II: mobility, share of population out of main center, distance from the airport; Controls III: hospital beds, nursing homes, share of families in poverty, hotel beds, household ownership, house price index, employment, teleworkability index, average temperature; Controls IV: Voter turnout. * ** $p < 0.01$, * * $p < 0.05$, * $p < 0.1$.

equivalent to 31% of the excess mortality of the period. For the South, a similar increase would reduce excess mortality by 0.188 between October and December 2021, imprecisely estimated (p.val. 0.156). In this case, there is no clear evidence of nonlinearities (Table C3 in the Online Appendix).

Last, we run a series of placebo tests which make explicit the role of the common trend hypothesis discussed and tested in appendix A. For each pre-pandemic year YY= 15.19 between 2015 and 2019, the period

Table 8
The effect of education on excess mortality. By Area. Placebos.

	(1) 2015	(2) 2016	(3) 2017	(4) 2018	(5) 2019
North					
Edu * Mar-May YY	-0.00774	-0.00116	0.00198	-0.00537	-0.00390
	(0.00637)	(0.00672)	(0.00560)	(0.00609)	(0.00671)
Centre-South					
Edu * Oct-Dec YY	-0.00726	-0.00631	-0.00524	-0.00164	-0.00569
	(0.00505)	(0.00492)	(0.00431)	(0.00485)	(0.00419)
Controls I*Period-Year Dummies	YES	YES	YES	YES	YES
Controls II*Period-Year Dummies	YES	YES	YES	YES	YES
Controls III*Period-Year Dummies	YES	YES	YES	YES	YES
Controls IV*Period-Year Dummies	YES	YES	YES	YES	YES

Note: We use model (1) to estimate the effect of education on excess mortality in year YY= 2015,…2019. In all cases excess mortality is defined as the difference between the monthly mortality rate among the 60 + in the period indicated in the rows in year YY and the average monthly mortality between 2012 and YY-1 in the same period. Robust standard errors clustered at the local labour system level in parentheses. Edu measures the share of individuals with at least upper secondary education. Controls I: average population age, population density, family size, share of migrants, male to female ratio, share of population older than 64; Controls II: mobility, share of population out of main center, distance from the airport; Controls III: hospital beds, nursing homes, share of families in poverty, hotel beds, household ownership, house price index, employment, teleworkability index, average temperature; Controls IV: Voter turnout. * ** $p < 0.01$, * * $p < 0.05$, * $p < 0.1$.

March-May for the North and the period October-December for the Centre-South, we test whether the excess mortality defined as $\Delta M_{iYYp} = M_{iYYp} - \overline{M}_{ip}^{12-(YY-1)}$ depends on municipality education. Intuitively, we test whether in the same periods where we find effects in 2020, we also find effects in the years not affected by the pandemic. Any significant pre-pandemic effect would cast doubts about model specification and identification. Reassuringly, results of Table 8 report point estimates which are always very small and never statistically significant.

7. Discussion

As the surveys collected during the pandemic lack either geographical or longitudinal detail, our ability to analyse the mechanisms behind the findings and test alternative explanations is limited. However, any plausible explanation must account for two facts: 1) education reduces the COVID-19-induced excess mortality; and 2) the effect of education emerges only in the early stage of the pandemic and disappears later on. Especially the failure to account for the second requirement allows to narrow down the set of alternatives.

To begin with, the effect of education cannot be ascribed to the correlation with, for instance, urbanization, mobility associated to tourism or commuting, population age, supply of healthcare facilities or social capital, because we directly control for many local characteristics capturing these dimensions (see Table 1).[18] However, the effect of education could depend on the correlation with omitted variables, such as the administrative quality of the local politicians. If better educated voters are more likely to elect better mayors, who adopted and enforced mobility restrictions and social distancing earlier than others, then municipalities with higher education would be a spared part of the mortality burden of the pandemic. This explanation, however, does not square well with the fact that, from the very beginning, the national government centralised all actions against COVID-19, and from March 9th 2020 a complete lockdown was imposed in the whole country.[19] Another possibility is that the effect of education could depend on its correlation with preference traits, such as risk and loss aversion (Jung, 2015; Dohmen et al., 2010; Benjamin et al., 2013), which are omitted from model (1). If, as suggested by Jung (2015), education and risk aversion are positively correlated, our results could simply indicate that the more risk-averse are more likely to adopt protective strategies. While we can't entirely exclude this possibility, it does not easily explain why the effect of education disappeared as the epidemic progressed.

The effect of education can depend on the fact that the more educated were more likely to work from home, while the less educated were more likely to work at their usual workplace. While Galasso and Foucault (2020) document that over 60% of college graduates worked from home during the first wave, they also document that over 60% of the less educated were home, mostly idle. During the following waves, the proportion of home working remained high and if anything, the less educated were more likely to return to their offices. Hence, the education gradient in the later waves should have been reinforced rather than weakened, even though protective measures were enforced in the workplaces. By using Google mobility data (defined at the province level), we try to directly test whether the reduction in mobility for work reasons varied differently in more and less educated provinces, by estimating a model akin to (1). The estimates do not reveal any statistically significant difference (Table C4 in the Online appendix).

We also check whether the effect of education can be ascribed to a possible differential reduction of road accidents during the lockdown, connected with the possibly lower mobility of the more educated. By using data from the Italian Automobile Club (ACI) from 2012 to 2020,[20] we regress the variations in road accidents between 2020 and the pre-pandemic period on municipal education and controls. We do not find evidence of any differential effect, as documented in Table C5 in the Online appendix.

Our conjecture is that the effect of education is genuine and has to do with the cognitive advantage accruing to the better educated people. Education affects people's cognition and their ability to understand the mechanics of virus transmission, the rationale of the proposed protective strategies, the credibility of information sources and, viceversa, the implausibility of "alternative" strategies (Freeman et al., 2020). Furthermore, the more educated individuals also have more access to information, because they read more and they are keener to follow the news and insights (Chan and Goldthorpe, 2007). Hence, being better able to discriminate between effective and ineffective protective strategies, the more educated manage to reduce the risk of infection (Reyna et al., 2009), and, if infected, the viral load (Goyal et al., 2021).[21]

This advantage is larger when information is confusing and inconsistent, as in this case the ability to discriminate information and judge its credibility turns to be more salient. In Appendix B, we document in some detail that information and public health communication, mixed and contradictory in the early phase of the pandemic, became more coherent later on. We also provide evidence of the surge of fake news during the first and partly during the second wave of the pandemic and the subsequent decline, by using data on Google search and Twitter. Besides the changing quality of information, a complementary cognitive process that adds to individual information set is that of learning from experience. As people learn from the outcomes of past decisions and infer what protective strategies are more effective, the initial advantage of the more educated narrows down. From this perspective, also the smaller effect of education observed in the Centre-South, the evidence of which is weaker compared to that found in the North, could be the result of the experience provided by the North (and other countries) during the first wave, which increased the awareness of Southern residents by the time COVID-19 reached their area, between October and December 2020. Hence, the interplay between education and the changing information context could explain the timing of the effects that we find in our analysis (protective effect in March-May 2020 in the North, and in October-December 2020 in the Centre-South, but no effect later on).[22]

To conclude, it is worth remarking that the effective protective strategies help prevent transmission to others, to the benefit also of the less educated. Therefore, in municipalities with higher average educational attainment, there is both a higher proportion of better-protected people, and a lower rate of transmission. Unfortunately, with municipality-level data these two dimensions cannot be disentangled, and our estimates reflect their combined contribution.

8. Conclusions

We have analysed the effect of education on excess mortality during the COVID-19 epidemic in Italy by exploiting detailed mortality data by municipality. Our results indicate that education played unambiguously a protective role in the North between March to May 2020 and there is some evidence of a protective role also in Centre-South between October and December 2020, although to a smaller extent. These are the periods when both areas experienced their first major COVID-19 outbreak. Differently, no effect of education was detected in later waves.

Our conjectural interpretation is that the protective effect depends on the cognitive advantage provided by education, while the timing of

[18] It is also unlikely that the intensity of touristic and business travel differed much between the two waves because mobility restrictions to long-range travel were in place in both waves.

[19] In the very first days of the pandemic, on February 27 2020, the mayor of Milan, one of the most educated municipalities, launched the campaign Milan does not stop (Milano non si ferma), around a video aiming to dispel worry and fear. Similar messages were launched on the same days by other mayors of important (and well educated) tows in Lombardy including Bergamo, Brescia. Twelve days later the government imposed the national lockdown. Two days later, on March 1st, the national government decided to close schools and universities and suspend sports events in Lombardy, Veneto, Emilia Romagna and three other provinces in Marche and Liguria. Lombardy was completely locked down by March 7th and by March 9th lockdown was extended to the entire country.

[20] Unfortunately, data for 2021 are unavailable.

[21] Another cognitive-related argument is suggested by the literature on innovation adoption. The more educated are more likely to be early adopters, because they have lower adoption costs and uncertainty about the value and the use of new technology (Wozniak, 1987; Riddell and Song, 2017). Particularly, Kämpfen and Maurer (2018) show that in Italy education had a strong causal effect on the adoption of new ICT among adults aged 50 + , and Lleras-Muney and Lichtenberg (2005) document that the more educated are more willing to use newly approved drugs in the US. As long as masks (of multiple types, cloth, surgical, or FFP2), face shields, sanitisers and tracing apps are considered new technologies, then the more educated should take them up first, and the less educated should follow later on.

[22] In the same direction goes the argument that the more educated are more likely to be early adopters of new technologies, as high-education municipalities would experience a more intense adoption of protective strategies during the first wave, and so, lower mortality, while, as time goes, the rate of adoption in all municipalities converges.

the effect may reflect the interplay between education on the one hand, and the coherence of public health messages, the reliability of COVID-19-related information, or a process of learning from experience on the other hand. In the early phase of the epidemic, information about preventive and protective measures against COVID-19 was confusing and contradictory. In this context, education helped discriminate between alternative sources and evaluate information reliability. Later on, information became more coherent and univocal, and/or people learned from their and their neighbours' experience what strategies were more effective, so that education ended up being less relevant. The changing effect of education as the pandemic progresses helps explaining why results are mixed in the small literature that analyses the effect of education in the spread and burden of COVID-19. Further research is needed to confirm our interpretation. To achieve firmer conclusions microdata including details on individual education, their adoption of protective behaviour and the evolution of the information set they have at the different stages of the pandemic is necessary.

Our findings add to the established evidence of the socio-economic gradient of health (Marmot, 2020), suggesting that better-educated people, in the midst of an unexpected crisis, are better able to cope. However, they also point out that the effect of education is not independent of the information that people obtain.

We draw two policy implications: first, supporting education, including adult education, might have important returns in the health domain and could help shelter people against possible new pandemics. Second, avoiding contradictory messages should be a primary concern for public health agencies and experts in order to prevent the rise of health inequalities associated with differential levels of education.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declarations of interest

None.

Data Availability

Data will be made available on request.

Acknowledgements

We are grateful Emanuele Caggiano for helpful discussions and having provided the data on tweets and retweets, to Giorgio Brunello, José Escarce, Vincenzo Galasso and Enrico Rettore for their helpful comments and suggestions.

Appendix A. – Common trend

Identification of the effect of education in model (1) depends on the assumption that mortality *levels* follow parallel trends between municipalities with different average education levels. Otherwise, the change in mortality between 2020 (2021) and the pre-pandemic era would vary between municipalities with different education regardless of the pandemic.

To test the parallel trend assumption, we state a model in terms of mortality level and exploit the detailed mortality data between 2012 and 2021. Hence for each period and area (North, Centre-South), we estimate continuous event study models specified as follows:

$$M_{ipt} = f_i + \alpha_{pt} + \beta_{pt} Edu_i + \delta_{pt} X_i + \varepsilon_{ipt} \quad for \quad each \quad p \in P \tag{2}$$

where P is the set of periods, namely January-February, March-May, June-September, October-December, t=2012.2021, M_{ipt} is the average monthly mortality rate in municipality i and period p of year t for the age group 60+, f_i are municipality fixed effects, Edu_i is the share of residents with at least upper secondary education in 2011, X_i is the set of control variables reported in Table 1. Systematic heterogeneity by municipality in the *level* of mortality is captured by municipality fixed effects. Common mortality trends are captured by α_{pt}. Systematic heterogeneity in the *trend* of mortality is parametrically accounted for by $\beta_{pt} Edu_i$ and $\delta_{pt} X_i$, which capture the differential period-year effects due to municipal education and other characteristics with respect to the common trend α_{pt}.[23] As usual, model (2) is identified up to a normalization, in our case $\alpha_{p2012} = 0$, $\beta_{p2012} = 0$ and $\delta_{p2012} = 0$, so that all parameters are to be interpreted as differential effects relative to 2012.

Under the parallel trend assumption, the effect of education in the COVID pandemic is the differential *change* in mortality in period p in 2020 (2021), with respect to the same period in 2012, between municipalities with different levels of education. Standard errors are clustered within local labour systems, a concept equivalent to that of commuting areas.

This model is akin to the seminal Card (1992), where the rise of the federal minimum wage, which applies to all US states in the same moment, is allowed to have a different effect on employment depending on the share of the teen population in each state.

A test of parallel trend amounts to test that the coefficients β_{pt} for $t = 2013, 2014 \ldots 2019$ are zero. To see why, take two sets of municipalities A and B with education equal to Edu_A and Edu_B respectively. For the sake of simplicity let us abstract from other characteristics X_i, or assume that they are equal between municipalities. According to model (2) expected mortality is $E(M_{Apt}|A, pt) = E(f_i|A, pt) + \alpha_{pt} + \beta_{pt} Edu_A$ and $E(M_{Bpt}|B, pt) = E(f_i|A, pt) + \alpha_{pt} + \beta_{pt} Edu_B$ and the mortality change in t=2013.2019 relative to 2012 (the last pre-pandemic year) are

$$E\left(M_{Apt}|A, pt\right) - E\left(M_{Apt}|A, p2012\right) = \alpha_{pt} + \beta_{pt} Edu_A$$

$$E\left(M_{Bpt}|B, pt\right) - E\left(M_{Bpt}|B, p2012\right) = \alpha_{pt} + \beta_{pt} Edu_B.$$

[23] Parameters β_{tp} and δ_{tp} are estimated by including full sets of interactions between period-year dummies and Edu_i, and between period-year dummies and each X_i, respectively. The omitted year dummy is the dummy referring to 2012.

The differential mortality change between the two sets of municipalities is $\beta_{pt}(Edu_A - Edu_B)$, which captures the differential mortality trend at year t in the pre-covid era. Under the parallel trend assumption, this quantity is zero, implying that $\beta_{pt} = 0$.

Estimates of model (2) are reported in Table A1. In no specification, we reject the hypothesis of parallel trends. This finding reinforces our confidence in the specification of model (1).[2] Estimates from model (2) are depicted in Figs. A1a and A1b for the North and the Centre-South respectively. Not only trends are parallel, they are also practically flat, confirming the visual evidence of Figures 2a and 2b. For Northern Italy, the anomaly of the period March-May 2020 compared to all other period-years is striking.

Table A1
Event Study 2012-2021. Education measured by share with at least upper secondary education.

	(1) North Jan-Feb	(2) North Mar-May	(3) North Jun-Sep	(4) North Oct-Dec	(5) Centre-South Jan-Feb	(6) Centre-South Mar-May	(7) Centre-South Jun-Sep	(8) Centre-South Oct-Dec
Edu x D13	0.0107	0.0114	0.0000	0.00889	-0.000160	0.00659	-0.00284	0.00758
	(0.0127)	(0.00812)	(0.00809)	(0.00756)	(0.00723)	(0.00726)	(0.00502)	(0.00626)
Edu x D14	0.00197	0.00332	0.00495	0.00726	0.000659	0.00661	0.000118	-0.00124
	(0.0119)	(0.00683)	(0.00698)	(0.00925)	(0.00766)	(0.00620)	(0.00465)	(0.00583)
Edu x D15	0.0108	-0.00318	-0.00708	0.0146*	0.00278	-9.39e-05	0.00617	-0.00515
	(0.0143)	(0.00832)	(0.00721)	(0.00770)	(0.00985)	(0.00541)	(0.00414)	(0.00608)
Edu x D16	0.0147	0.00159	-0.00225	0.00515	-0.00687	0.00268	-0.00259	-0.00601
	(0.0139)	(0.00753)	(0.00814)	(0.00904)	(0.00834)	(0.00672)	(0.00498)	(0.00553)
Edu x D17	0.0162	0.00335	0.00233	0.0126	0.00265	0.00115	-0.00716	-0.00621
	(0.0113)	(0.00712)	(0.00676)	(0.00784)	(0.00848)	(0.00616)	(0.00515)	(0.00521)
Edu x D18	0.00557	-0.00417	-0.00550	0.00916	-0.00430	-0.00379	-0.00933*	-0.00347
	(0.0128)	(0.00805)	(0.00739)	(0.00745)	(0.00839)	(0.00695)	(0.00503)	(0.00586)
Edu x D19	0.00859	0.00331	0.00191	0.0112	0.00161	-0.00874	-0.00690	-0.00777
	(0.0109)	(0.00797)	(0.00718)	(0.00737)	(0.00829)	(0.00672)	(0.00467)	(0.00607)
Edu x D20	0.0112	-0.0421**	0.000414	0.0147*	-0.00558	0.00149	-0.00182	-0.0109**
	(0.0141)	(0.0165)	(0.00678)	(0.00858)	(0.00826)	(0.00713)	(0.00596)	(0.00522)
Edu x D21	0.0152	-0.00604	-0.00158		-0.00499	-0.00416	-0.00534	
	(0.0136)	(0.00884)	(0.00723)		(0.00805)	(0.00793)	(0.00541)	
Observations	42,350	42,350	42,350	38,095	33,030	33,030	33,030	29,727
R-squared	0.188	0.255	0.216	0.231	0.179	0.197	0.202	0.200
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Municipalities FE	YES	YES	YES	YES	YES	YES	YES	YES
Controls x Year Dummies included	YES	YES	YES	YES	YES	YES	YES	YES
Average Morality Rate (12-19)	3.534	3.006	2.758	3.047	3.833	3.300	2.944	3.212
Excess Mortality Rate 20	-0.396	1.783	-0.041	1.151	-0.389	0.096	0.004	0.562
Excess Mortality Rate 21	-0.030	0.250	0.200		-0.017	0.273	0.199	
Observations	42,350	42,350	42,350	38,095	33,030	33,030	33,030	29,727

Note: Years: 2012-2021. Robust standard errors clustered at the local labour system level in parentheses. The dependent variable is the monthly mortality rate for the age group aged 60+ in the period reported in each column head. Edu measures the share of individuals with at least upper secondary education, and Edu x Dyy are interactions between Edu and year dummies All regressions include interactions between all controls listed in Table 1 and Dyy, year dummies and municipalities fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

In discussing model (1), we named the dependent variable excess mortality, meaning abnormal mortality relative to a usual constant level. This label turns out to be fully correct given the evidence of flat mortality trends.

We conclude this section by showing that model (1) immediately derives from model (2).

First, we average model (2) between 2012 and 2019 for each municipality and obtain

$$\overline{M}_{ip}^{12-19} = f_i + \overline{\alpha}_p + \overline{\beta}_p Edu_i + \overline{\delta}_p X_i + \overline{\varepsilon}_{ip} \tag{3}$$

Next for t=2020,2021 and all p we subtract (3) from (2) and obtain

[2] At the beginning of the pandemic, public health messages about the mechanics of COVID-19 transmission and the protective strategies were quite discordant. Reputed doctors claimed that COVID-19 was not worse than a normal flu and that stricter preventive measures were unnecessary. Emblematic of these contradictions was the discussion about the opportunity of wearing face masks. We also document, by exploiting data from Google Trends and Twitter, that the diffusion of fake news regarding supposedly miraculous preventive strategies, such as taking Vitamin C or disinfecting the throat with bleach, reached its peak between March and May 2020 and then declined.

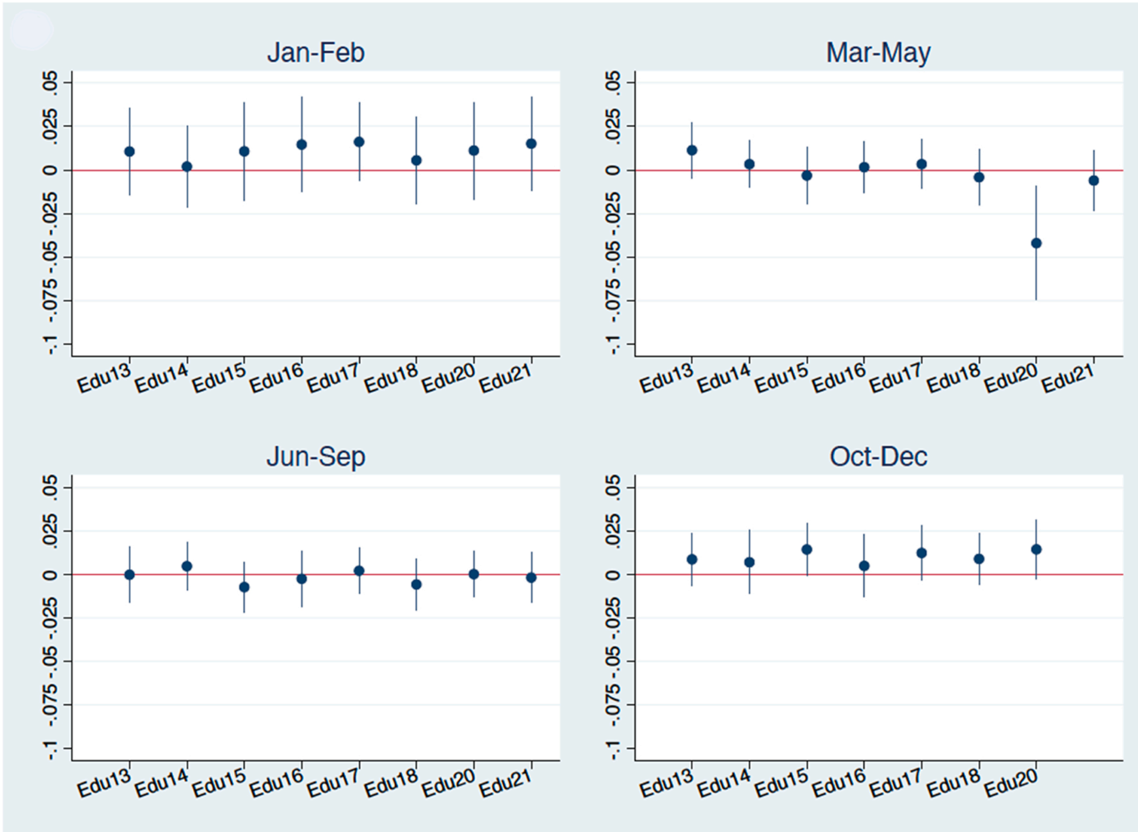


Fig. A1.a. Estimates of the differential effect of education between 2012 and 2021 (model 2), North Italy. *Note: The figures show the coefficients of the interaction terms between Education and the year dummies from Model 2 by stage of the pandemic. The omitted year is 2012. Source: ISTAT.*

Fig. A1.b. Estimates of the differential effect of education between 2012 and 2021 (model 2), Centre-South Italy. *Note: The figures show the coefficients of the interaction terms between Education and the year dummies from Model 2 by stage of the pandemic. The omitted year is 2012. Source: ISTAT.*

$$M_{ipt} - \overline{M}_{ip}^{12-19} = (\alpha_{pt} - \overline{\alpha}_p) + (\beta_{pt} - \overline{\beta}_p)Edu_i + (\delta_{pt} - \overline{\delta}_p)X_i + (\varepsilon_{ipt} - \overline{\varepsilon}_{ip}) \quad (4)$$

which is equivalent to (1). Note that this differencing allows to cancel municipality fixed effects implying that model (1) is robust to any time invariant unobserved heterogeneity in mortality levels across municipalities.

The parameter attracted by Edu_i is $(\beta_{pt} - \overline{\beta}_p)$, that is the differential effect of education in 2020 (resp. 2021) compared with the pre-pandemic effect $\overline{\beta}_p$. The parallel trend hypothesis and the normalization required in model (2) imply that $\overline{\beta}_p = 0$ and that, absent the pandemic, $(\beta_{pt} - \overline{\beta}_p) = 0$.

Appendix B. – COVID information and public health communication during the pandemic

At the beginning of the COVID-19 pandemic, information was incomplete, contradictory and ridden with fake news. For instance, it was deemed unlikely that healthy people could spread the virus and hence testing was limited and restricted to the symptomatic. There was also quite some debate about the protective measures to be adopted. Very telling of this confusion was the communication about the use of face masks.[3] WHO guidelines issued on January 29th maintained that masks should be used only by the healthcare personnel and not by the general public. This was also the official position of Italian authorities, despite many scientists had opposite views, and some regions started issuing contrasting messages. For instance, on the 3rd of April, the head of the Italian Civil Protection Agency, Angelo Borrelli, stated that he was not wearing masks, although he kept the social distancing of one meter. The same day, the Lombardy region issued a regulation that made the use of masks compulsory outdoors. On the day after, April 4th, the Chief Public Health Officer, Franco Locatelli,[4] declared that there was no firm evidence of the effectiveness of masks, while Andrea Crisanti, the researcher who first proved the role of the asymptomatic in spreading the virus, simultaneously claimed that “masks are key: better use them also indoor”.

In June (June 5th) the WHO updated its guidelines and stated that “[.] governments should encourage the general public to wear masks in specific situations and settings (WHO, 2020b, p.6)” and only on December 1st, it stated that “the general public should wear a non-medical mask in indoor (e.g. shops, shared workplaces, schools [.]) or outdoor settings where physical distancing of at least 1 meter cannot be maintained. (WHO 2020c, p.1)”.

In Italy coherence between public health messaging was also complicated by the regional organization of healthcare. All Italian regions have important competencies regarding health and prevention, which should be coordinated with the central bodies at the ministry of public health. Especially during the first phases of the pandemic, however, coordination was difficult and regions took autonomous initiatives (Antonini et al. 2020; Berardi et al. 2020). Regional decisions were reported by the national media, reached the general public (Castriota et al. 2020) and increased confusion.

During the early phases of the epidemic, there was also a sustained circulation of fake news. For instance, in social networks, many people claimed that the intake of Vitamine C and the practice of gargles with bleach were effective preventive strategies.[5] We used Google Trends and Twitter to document the popularity trend of these unscientific treatments.

In Fig. B1.a and B1.b, we report the popularity index of the words “Vitamin C” and “Bleach” from Google Trends, between 2012 and 2021, by period-year and area. Google trends re-scales the time series of the number of searchers on the Google engine by month and region, assigns 100 to the date of the highest search count, and a proportional number to all other dates. It also provides an analogous index across regions at a given date. After noticing that for all regions the peak of the time series occurred in March 2020, we used the cross-section to rescale all region-specific time series and make them comparable. Finally, we aggregate by period-year and area. For both “Vitamin C” and “Bleach” the popularity was highest between March and May 2020. Popularity fast declined after March-May 2020 although it remained slightly higher in the Centre-South than in the North.

Turning to Twitter, we took all tweets which contain the phrase “Vitamin C” or “Bleach”, which were re-tweeted at least once. The latter qualification is aimed to isolate tweets that had at least some impact. We sum the number of tweets and retweets in each period-year in Italy.[6] Results are reported in Fig. B2.a and B2.b. The pattern is analogous to that obtained from Google Trends, with the highest number of tweets including the words “Vitamin C” and “Bleach” recorded between March and May 2020. While for “Vitamin C” there were some spikes before and after the peak, for “Bleach” there was a unique spike and the number of tweets almost reverted to pre-pandemic levels soon after the peak.

Contradictory information and unscientific claims increase uncertainty about the effectiveness of each prevention strategy, causing underadoption (Webster et al. 2020). Nonetheless, the more educated are better able to discriminate and elicit information coming from reliable sources, they are more likely to dismiss unscientific claims and eventually, uncertainty is lower to them.

[3] By excess mortality rate we mean the absolute difference between the mortality rate among the over-60 recorded in 2020 and that recorded on average between 2012 and 2019. Mortality rates are computed as the number of deaths in the age group 60 + per 1000 residents aged 60 + , per month.

[4] We define Mits = (deaths in period s of year t among the population aged 60 + in municipality i)/(population aged 60 + in year t in municipality i) x (1000/Ns), where Ns is the number of months in period s.

[5] The education system is homogenous in Italy, and there are no marked differences in education attainment across country areas. In the South, the share of residents with at least upper secondary education in 2011 was 41 percent, compared to 39 percent in the more economically advanced North. Only the Centre somehow differs with its 46 percent. However, school quality is not even and there is evidence of a more marked and persistent heterogeneity in the South than elsewhere (INVALSI, 2019).

[6] The proxy for education is the municipal share of residents with at least upper secondary education. The figures also provide the map for the same period in 2019. For the pre-pandemic year, there is no evidence of a correlation between municipal education and excess mortality.

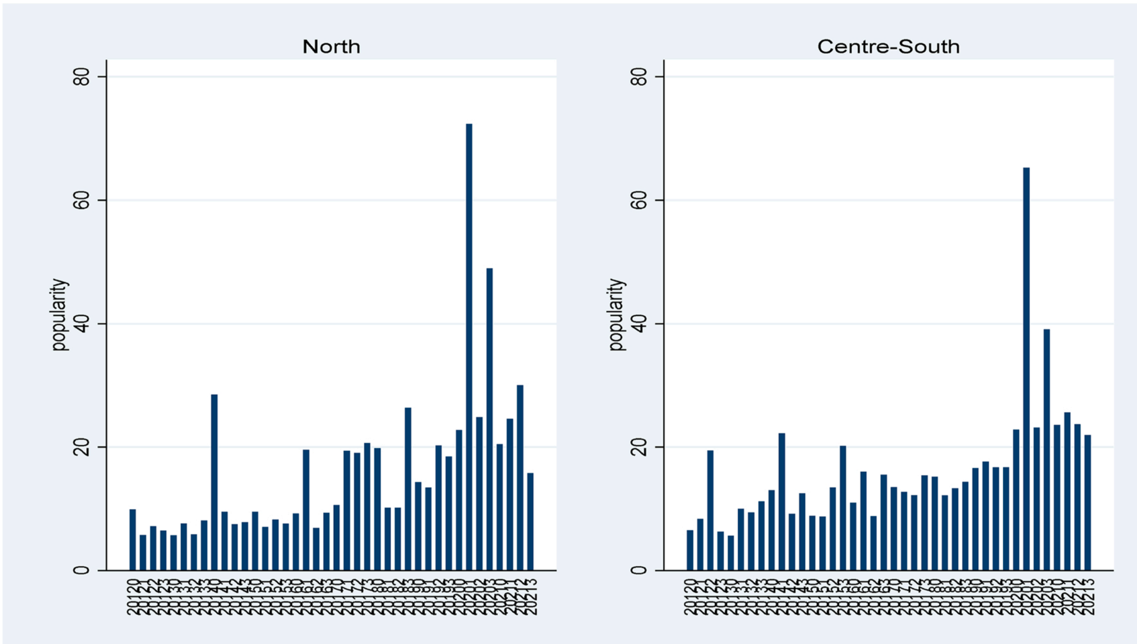


Fig. B1.a. Popularity of search including "Vitamin C" from Google Trends. By Area.

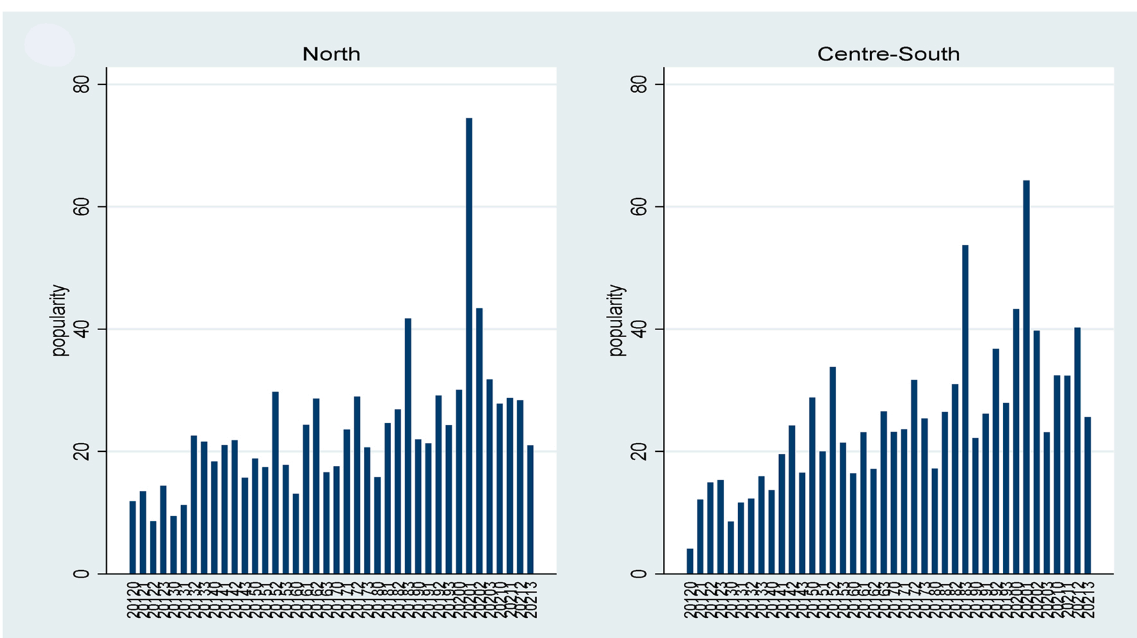


Fig. B1.b. : Popularity of search including "Bleach" from Google Trends. By Area. *Note: The figures show the average searches (on a scale from 0-100) for "Vitamine C" and "Bleach" ("Vitamina C" and "Candeggina" in Italian) in each year and month by area (North and Centre-South). Source: Google Trends.*

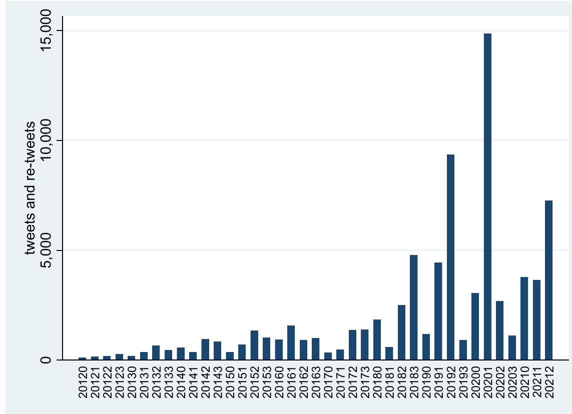


Fig. B2.a. : Number of tweets and retweets including "Vitamin C".

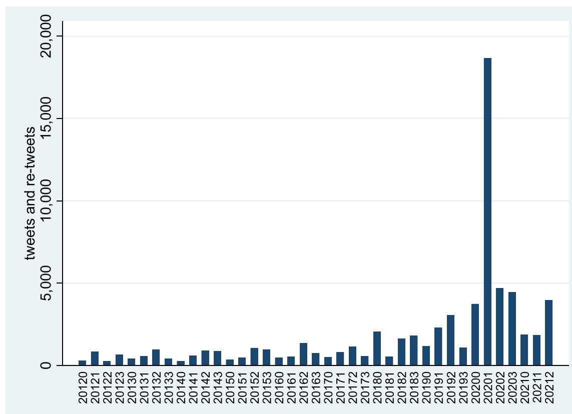


Fig. B2.b. Number of tweets and retweets including "Bleach". *Note: The figures show the average number of Tweets and reTweets including "Vitamin C" and "Bleach" ("Vitamina C" and "Candeggina" in Italian) in each year and month. Source: Twitter.*

Appendix C. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ehb.2022.101194.

References

Adams, S.J., 2002. Educational attainment and health: evidence from a sample of older adults. Educ. Econ. 10 (1), 97–109.

Alacevich, C., Cavalli, N., Giuntella, O., Lagravinese, R., Moscone, F., Nicodemo, C., 2021. The presence of care homes and excess deaths during the COVID-19 pandemic: evidence from Italy. Health Econ. 30, 1703–1710.

Almagro, M., Orane-Hutchinson, A., 2020. JUE insight: the determinants of the differential exposure to COVID-19 in New York city and their evolution over time. J. Urban Econ., 103293

Antonini, M., Berardi, C., Paolucci, F., 2020. Le Regioni in ordine sparso nella lotta al COVID. Quotid. Sanità 2020.

Aparicio Fenoll, A., Grossbard, S., 2020. Intergenerational residence patterns and Covid-19 fatalities in the EU and the US. Econ. Hum. Biol. 39, 100934.

Arendt, J.N., 2005. Does education cause better health? A panel data analysis using school reforms for identification. Econ. Educ. Rev. 24, 149–160.

Armillei, F., Francesco, F., Fletcher, T., 2021. "Did COVID-19 hit harder in peripheral areas? The case of Italian municipalities". Econ. Hum. Biol. 42, 101018.

Barrios, J.M., Benmelech, E., Hochberg, Y.V., Sapienza, P., Zingales, L., 2021. Civic capital and social distancing during the COVID-19 pandemic. J. Public Econ. 193, 104310.

Bartscher, A.K., Seitz, S., Siegloch, S., Slotwinski, M., Wehrhöfer, N., 2021. Social capital and the spread of covid-19: insights from European countries. J. Health Econ. 80, 102531.

Beaney, T., Clarke, J.M., Jain, V., Kataria Golestaneh, A., Lyons, G., Salman, D., Majeed, A., 2020. Excess mortality: the gold standard in measuring the impact of COVID-19 worldwide. J. R. Soc. Med. 113 (9), 329–334.

Benjamin, D.J., Brown, S.A., Shapiro, J.M., 2013. Who is 'behavioral'? Cognitive ability and anomalous preferences. J. Eur. Econ. Assoc. 11 (6), 1231–1255.

Berardi, C., Antonini, M., Genie, M.G., Cotugno, G., Lanteri, A., Melia, A., Paolucci, F., 2020. The COVID-19 pandemic in Italy: policy and technology impact on health and non-health outcomes. Health Policy Technol. 9 (4), 454–487.

Bish, A., Michie, S., 2010. Demographic and attitudinal determinants of protective behaviours during a pandemic: a review. Br. J. Health Psychol. 15, 797–824.

Bisin, A. and Moro, A. (2020) Learning Epidemiology By Doing: The Empirical Implications Of A Spatial-Sir Model With Behavioral Responses. NBER working paper n. 27590.

Borgonovi, F., Andrieu, E., Subramanian, S.V., 2021. The evolution of the association between community level social capital and COVID-19 deaths and hospitalizations in the United States. Soc. Sci. Med. 278, 113948.

Borjas, G.J. (2020). Demographic determinants of testing incidence and COVID-19 infections in New York City neighborhoods. NBER working paper n. 26952.

Brandily, P., Brébion, C., Briole, S., Khoury, L., 2020. A poorly understood disease? The unequal distribution of excess mortality due to COVID-19 across French municipalities. NHH Dept. Econ. Discuss. 15.

Brodeur, A., Grigoryeva, I., Kattan, L., 2020. Stay-at-home orders, social distancing and trust. J. Popul. Econ. 34, 1321–1354.

Callaway, B. and Li, T. (2021) Policy Evaluation during a Pandemic, mimeo.

Card, D., 1992. Using regional variation in wages to measure the effects of the federal minimum wage. Ind. Labor Relat. Rev. 46 (1), 22–37.

Castriota, S., Delmastro, M., & Tonin, M. (2020). National or local? The demand for news in Italy during COVID-19. CESifo Working Paper n. 8699.

Chan, T.W., Goldthorpe, J.H., 2007. Social status and newspaper readership. Am. J. Sociol. 112 (4), 1095–1134.

Charoenwong, B., Kwan, A., Pursiainen, V., 2020. Social connections with COVID-19–affected areas increase compliance with mobility restrictions. Sci. Adv. 6 (47), eabc3054.
Cheng, Y., Ma, N., Witt, C., Rapp, S., Wild, P.H., Andreae, M.O., Pöschl, U., Su, Hang, 2021. Face masks effectively limit the probability of SARS-CoV-2 transmission. Science 372, 1439–1443.
Conti, G., Heckman, J., Urzua, S., 2010. The education-health gradient. Am. Econ. Rev.: Pap. Proc. 100, 234–238.
Crighton, E.J., Elliott, S.J., Moineddin, R., Kanaroglou, P., Upshur, R., 2007. A spatial analysis of the determinants of pneumonia and influenza hospitalizations in Ontario (1992–2001). Soc. Sci. Med. 64 (8), 1636–1650.
Cutler, D.M., Lleras-Muney, A., 2010. Understanding differences in health behaviors by education. J. Health Econ. 29, 1–28.
Daon, Y., Thompson, R.N., Obolski, U., 2020. Estimating COVID-19 outbreak risk through air travel. J. Travel Med. 1–8. https://doi.org/10.1093/jtm/taaa093.
Desmet, K., Wacziarg, R., 2021. JUE Insight: Understanding spatial variation in COVID-19 across the United States. J. Urban Econ., 103332
Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? Am. Econ. Rev. 100 (3), 1238–1260.
Durante, R., Guiso, L., Gulino, G., 2021. Asocial capital: Civic culture and social distancing during COVID-19. J. Public Econ. 194, 104342.
Freeman, D., Waite, F., Rosebrock, L., Petit, A., Causier, C., East, A., Lambe, S., 2020. Coronavirus conspiracy beliefs, mistrust, and compliance with government guidelines in England. Psychol. Med. 52 (2), 251–263.
Galasso, V. and Foucault, M. (2020). Working during COVID-19: Cross-country evidence from real-time survey data. OECD Social, Employment and Migration Working Papers n. 246.
Galasso, V., Pons, V., Profeta, P., Becher, M., Brouard, S., Foucault, M., 2020. Gender differences in COVID-19 attitudes and behavior: panel evidence from eight countries. Proc. Natl. Acad. Sci. 117 (44), 27285–27291.
Gallotti, R., Valle, F., Castaldo, N., Sacco, P., De Domenico, M., 2020. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. Nat. Hum. Behav. 4, 1285–1293.
Ginsburgh, V., Magerman, G., Natali, I., 2021. COVID-19 and the role of inequality in French regional departments. *The*. Eur. J. Health Econ. 22 (2), 311–327.
Glaeser, E.L., Gorback, C.S., & Redding, S.J. (2020). How much does COVID-19 increase with mobility? Evidence from New York and four other US cities. NBER working paper n. 27519.
González-Val, R., Marcén, M., 2021. Mass gathering events and the spread of infectious diseases: evidence from the early growth phase of COVID-19. Econ. Hum. Biol. 46, 101140.
Goyal, A., Reeves, D.B., Thakkar, N., Famulare, M., Cardozo-Ojeda, E.F., Mayer, B.T., Schiffer Marmot, J.T., 2020. Slight reduction in SARS-CoV-2 exposure viral load due to masking results in a significant reduction in transmission with widespread implementation. Nat. Sci. Rep. 11, 11838.
Guiso, L., Sapienza, P., Zingales, L., 2004. The role of social capital in financial development. Am. Econ. Rev. 94 (3), 526–556.
Healy, J.D., 2003. Excess winter mortality in Europe: a cross country analysis identifying key risk factors. J. Epidemiol. Community Health 57, 784–789.
Hong, K., Savelyev, P.A., Tan, K.T.K., 2020. Understanding the Mechanisms Linking College Education with Longevity. J. Hum. Cap. 14 (3), 371–400.
INVALSI (2019). Rapporto Prove Invalsi 2019. Invalsi, Rome.
Irons, N.J., Raftery, A.E., 2021. Estimating SARS-CoV-2 infections from deaths, confirmed cases, tests, and random surveys. PNAS 118 (31), e2103272118.
Jung, S., 2015. Does education affect risk aversion? Evidence from the British education reform. Appl. Econ. 47 (28), 2924–2938.
Kämpfen, F., Maurer, J., 2018. Does education help "old dogs" learn "new tricks"? The lasting impact of early-life education on technology use among older adults. Res. Policy 47, 1125–1132.
Lerchl, A., 1998. Changes in the seasonality of mortality in Germany from 1946 to 1995: the role of temperature. Int. J. Biometeorol. 42, 84–88.
Lleras-Muney, A., Lichtenberg, F.R., 2005. Are the more educated more likely to use new drugs? Ann. D. 'Économie Et. De. Stat. 79/80, 671–696.
Loomba, S., de Figueiredo, A., Piatek, S.J., de Graaf, K., Larson, H.J., 2021. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nat. Hum. Behav. 5, 337–348.
Marmot, M., 2020. Health equity in England: the Marmot review 10 years on. BMJ 368, m693.
Marschner, I.C., 2021. Estimating age-specific COVID-19 fatality risk and time to death by comparing population diagnosis and death patterns: Australian data. BMC Med. Res. Methodol. 21, 126.
Mecenas, P., Travassos da Rosa Moreira Bastos, R., Vallinoto, A.C.R., Normando, D., 2020. Effects of temperature and humidity on the spread of COVID-19: a systematic review. PLOS ONE 15 (9), e0238339.
Mills, S.D., Charlton, S.R., and Rowley, B.M. (2016). "Evaluation of Education and Other Influential Factors on the Perceptions of Influenza Vaccinations. Journal of the Arkansas Academy of Science. 70, Article 28.
Nagata, J.M., Hernández-Ramos, I., Kurup, A.S., et al., 2013. Social determinants of health and seasonal influenza vaccination in adults ≥65 years: a systematic review of qualitative and quantitative data. BMC Public Health 13, 388.
Nivette, A., Ribeaud, D., Murray, A., Steinhoff, A., Bechtiger, L., Hepp, U., Eisner, M., 2021. Non-compliance with COVID-19-related public health measures among young adults in Switzerland: Insights from a longitudinal cohort study. Soc. Sci. Med. 268, 113370.
Notari,A. and Torrieri, G. (2021) COVID-19 transmission risk factors. mimeo.
Ornaghi, C., & Tonin, M. (2018). Water Tariffs and Consumers' Inaction. CESifo Working Paper Series n. 6990.
Reyna, V.F., Nelson, W.L., Han, P.K., Dieckmann, N.F., 2009. How numeracy influences risk comprehension and medical decision making. Psychol. Bull. 135 (6), 943.
Riddell, W.C., Song, X., 2017. The role of education in technology use and adoption: evidence from the canadian workplace and employee survey. ILR Rev. 70 (5), 1219–1253.
Rolfes, M.A., Foppa, I.M., Garg, S., et al., 2018. Annual estimates of the burden of seasonal influenza in the United States: a tool for strengthening influenza surveillance and preparedness. Influenza Other Respir. Virus 12, 132–137.
Sanmarchi, F., Golinelli, D., Lenzi, J., Esposito, F., Capodici, A., Reno, C., Gibertoni, D., 2021. Exploring the gap between excess mortality and COVID-19 deaths in 67 countries. JAMA Netw. Open 4 (7), e2117359.
Sostero, M., Milasi, S., Hurley, J., Fernández-Macías, E., and Bisello, M. (2020) Teleworkability and the COVID-19 crisis: a new digital divide? JRC working papers series on labour, education and technology, n. 2020/05.
Tabellini, G., 2010. Culture and institutions: economic development in the regions of Europe. J. Eur. Econ. Assoc. 8 (4), 677–716.
Tsai, J., Traub, E., Aoki, K., Oyong, K., Sato, H., Rizik-Baer, D., Gounder, P., 2021. Incidentally detected SARS-COV-2 among hospitalized patients in los angeles county, August to October 2020. J. Hosp. Med. 16 (8), 480–483.
Verwimp, P. (2020). The Spread of COVID-19 in Belgium: a Municipality-Level Analysis. ECARES Working Paper n.25.
Ward, T., Johnsen, A., 2021. Understanding an evolving pandemic: an analysis of the clinical time delay distributions of COVID-19 in the United Kingdom. PLoS ONE 16 (10), e0257978.
Webster, R.K., Brooks, S.K., Smith, L.E., Woodland, L., Wessely, S., Rubin, G.J., 2020. How to improve adherence with quarantine: rapid review of the evidence. Public Health 182, 163–169.
WHO (2020a). "Coronavirus disease (COVID-19): How is it transmitted? 2020." URL: ⟨https://www.who.int/news-room/qa-detail/qa-how-is-covid-19-transmitted⟩.
WHO (2020b). Advice on the use of masks in the context of COVID-19: interim guidance, 5 June 2020.World Health Organization.
WHO (2020c). Mask use in the context of COVID-19: interim guidance, 1 December 2020. World Health Organization.
WHO (2021) Infodemic management, World Health Organization.
Wiliński, A., Kupracz, Ł., Senejko, A., Chrząstek, G., 2022. COVID-19: average time from infection to death in Poland, USA, India and Germany. Qual. Quant. https://doi.org/10.1007/s11135-022-01340-w.
Wozniak, G.D., 1987. Human capital, information, and the early adoption of new technology. J. Hum. Resour. 22 (1), 101–112.
Wright, L., Steptoe, A., Fancourt, D., 2021. Patterns of compliance with COVID-19 preventive behaviours: a latent class analysis of 20,000 UK adults. medRxiv Prepr. https://doi.org/10.1101/2021.03.16.21253717.
Wu, S.L., Mertens, A.N., Crider, Y.S., Nguyen, A., Pokpongkiat, N.N., Djajadi, S., Seth, A., Hsiang, M.S., Colford, J.M., Reingold, A., 2020. Substantial underestimation of SARS-CoV-2 infection in the United States. Nat. Commun. 11, 4507.
Yang, K., Pierri, F., Hui, P., Axelrod, D., Torres-Lugo, C., Bryden, J., Menczer, F., 2021. The COVID-19 infodemic: Twitter versus Facebook. *Big Data Soc.*, Jan. 1–16.
Zarocostas, J. (2020) How to fight an infodemic. The Lancet 395 Feb.29, 2020.
Zhang, Y.S.D., Leslie, H.Y., Sharafaddin-Zadeh, Y., Noels, K., Lou, N.M., 2021. Public health messages about face masks early in the COVID-19 pandemic: perceptions of and impacts on Canadians. J. Community Health 46, 903–912.